72182

STIC-Biotech/ChemLib

From: Turner, Sharon
Sent: Wednesday, July 31, 2002 9:12 AM
To: STIC-Biotech/ChemLib
Subject: 08948131

Please search CDB and Interference files

08948131

Oligomer search SEQ ID NO:1

Thanks,
Sharon L. Turner, Ph.D.
USPTO
CM1-10B09
Mailroom 10C01
Biotechnology GAU 1647
(703) 308-0056

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

7-26-01 gh

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 7/31/02
Date Completed: 7/31/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 10
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:03:45 ; Search time 30.19 Seconds
(without alignments)
110.375 Million cell updates/sec

Title: US-08-948-131-1
Perfect score: 30
Sequence: 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	30	20	AAY09349	Human RAGE V-domai
2	30	100.0	30	21	AAY52134	Human Receptor to
3	30	100.0	318	18	AAW44200	Human mature recep
4	30	100.0	318	18	AAW33754	Human RAGE polypep
5	30	100.0	332	21	AAY52130	Human Receptor to
6	30	100.0	340	18	AAW44199	Human soluble rece
7	30	100.0	340	18	AAW33753	Human RAGE polypep
8	30	100.0	404	22	AAB81925	Extracorporeal cir
9	16	53.3	16	18	AAW44214	Human soluble RAGE
10	16	53.3	16	18	AAW33768	Human RAGE polypep
11	15	50.0	15	18	AAW44208	Human soluble RAGE
12	15	50.0	15	18	AAW33762	Human RAGE polypep
13	13	43.3	30	20	AAY09350	Mouse RAGE V-domai
14	11	36.7	30	20	AAY09351	Rat RAGE V-domain
15	10	33.3	10	18	AAW44209	Human soluble RAGE
16	10	33.3	10	18	AAW44201	Human soluble RAGE
17	10	33.3	10	18	AAW33763	Human RAGE polypep
18	10	33.3	10	18	AAW33755	Human RAGE polypep
19	10	33.3	10	20	AAY09353	Human RAGE V-domai
20	10	33.3	10	21	AAY52135	Human Receptor to
21	10	33.3	30	20	AAY09352	Bovine RAGE V-doma
22	9	30.0	30	18	AAW23337	N-terminal seqeunc
23	7	23.3	178	22	AAB36499	Acinetobacter sp.
24	7	23.3	237	22	ABB67569	Drosophila melanog
25	7	23.3	742	22	ABB59085	Drosophila melanog
26	6	20.0	10	18	AAW27362	Residues 138-147 o
27	6	20.0	10	20	AAW92647	Human HAI-1 peptid
28	6	20.0	39	22	AAO09419	Human polypeptide
29	6	20.0	47	22	AAO11787	Human polypeptide
30	6	20.0	49	22	ABB42915	Peptide #10421 enc
31	6	20.0	49	22	ABB26166	Protein #8165 enco
32	6	20.0	49	22	AAM63816	Human brain expres
33	6	20.0	49	22	AAM76630	Human bone marrow
34	6	20.0	49	22	AAM36735	Peptide #10772 enc
35	6	20.0	61	22	ABB15264	Human nervous syst
36	6	20.0	78	22	AAU52300	Propionibacterium
37	6	20.0	86	22	AAO03946	Human polypeptide
38	6	20.0	92	22	AAO08061	Human polypeptide
39	6	20.0	99	22	AAU55509	Propionibacterium
40	6	20.0	105	21	AAG57767	Arabidopsis thalia
41	6	20.0	114	22	AAM91827	Human immune/haema
42	6	20.0	116	21	AAB41577	Human ORFX ORF1341
43	6	20.0	116	22	AAO05586	Human polypeptide
44	6	20.0	139	22	AAU66290	Propionibacterium
45	6	20.0	139	22	AAO08263	Human polypeptide

ALIGNMENTS

RESULT 1
AAY09349
ID AAY09349 standard; peptide; 30 AA.
XX
AC AAY09349;
XX
DT 09-JUL-1999 (first entry)
XX
DE Human RAGE V-domain peptide SEQ ID NO:1.
XX
KW RAGE; V-domain; receptor for advanced glycation endproduct;
KW ligand binding site; amyloid-beta; Alzheimer's disease; Down's syndrome;
KW senility; renal failure; hyperlipidaemic atherosclerosis; dementia;
KW neuronal cytotoxicity; head trauma; amyotrophic lateral sclerosis;
KW multiple sclerosis; amyloidosis; autoimmune disease; inflammation;
KW tumour; cancer; male impotence; wound healing; periodontal disease;
KW neuropathy; retinopathy; nephropathy; neuronal degeneration.
XX
OS Homo sapiens.
XX
PN WO9918987-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21346.
XX
PR 09-OCT-1997; 97US-0948131.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Lamster I, Schmidt AM, Stern D, Yan SD;
XX
DR WPI; 1999-277439/23.
XX

PT New peptides based on an advanced glycation end product receptor are
PT useful for treating Alzheimer's disease and Down's syndrome
XX
PS Claim 2; Page 78; 101pp; English.
XX
CC The present invention describes novel isolated peptides (I) having an
CC amino acid sequence corresponding to an amino acid sequence of a
CC V-domain of a receptor for an advanced glycation end product (RAGE).
CC Also described are methods for: (1) inhibiting an amyloid-beta peptide
CC (ABP) interaction with a receptor for RAGE when the receptor is on the
CC surface of a cell; (2) inhibiting degeneration of a neuronal cell; (3)
CC inhibiting formation of an ABP fibril on a cell; (4) inhibiting
CC extracellular assembly of an ABP into a fibril; (5) inhibiting
CC aggregation of ABP on the surface of a cell; (6) inhibiting infiltration
CC of a microglial cell into senile plaques; (7) inhibiting activation of a
CC microglial cell by an ABP; (8) treating a subject with a condition
CC associated with an interaction of an ABP with a receptor for RAGE on a
CC cell; (9) evaluating the ability of an agent to inhibit binding of an
CC ABP with a V-domain of a receptor for RAGE on the surface of a cell; (10)
CC inhibiting activation of a NF-kappaB gene in a cell; (11) inhibiting
CC periodontal disease in a subject; (12) inhibiting an RAGE's interaction
CC with a receptor for RAGE when the receptor is on the surface of a cell;
CC and (13) treating a subject with a condition associated with an
CC interaction of an RAGE with a receptor for RAGE on a cell. The methods
CC can be used for treating conditions associated with an interaction of an
CC ABP or an RAGE with a receptor for RAGE, e.g. diabetes, Alzheimer's
CC disease, senility, renal failure, hyperlipidaemic atherosclerosis,
CC neuronal cytotoxicity, Down's syndrome, dementia associated with head
CC trauma, amyotrophic lateral sclerosis, multiple sclerosis, amyloidosis,
CC an autoimmune disease, inflammation, a tumour, cancer, male impotence,
CC wound healing, periodontal disease, neuropathy, retinopathy, nephropathy
CC or neuronal degeneration.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.8e-25;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||||||||
Db 1 aqnitarigeplvlkckgapkkppqrlewk 30

RESULT 2
AAY52134
ID AAY52134 standard; protein; 30 AA.
XX
AC AAY52134;
XX
DT 28-JAN-2000 (first entry)
XX
DE Human Receptor to AGE (RAGE) amino acid sequence fragment #3.
XX
KW Soluble receptor for advanced glycation endproducts; RAGE; tumour;
KW invasion; metastasis; amphoterin; neuron; inhibit; therapy.
XX
OS Homo sapiens.
XX
PN WO9954485-A1.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US08427.
XX
PR 17-APR-1998; 98US-0062365.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Schmidt AM, Stern D;
XX
DR WPI; 2000-013260/01.
XX
PT Inhibiting tumour invasion or spreading by administration of soluble
PT receptor for advanced glycation endproducts -
XX
PS Claim 29; Page 62; 88pp; English.
XX
CC This is the amino acid sequence of a fragment of the human soluble
CC Receptor for Advanced Glycation Endproducts (RAGE). RAGE interacts with
CC a range of physiologically and pathophysiologically relevant ligands
CC when considering tumour invasion. In normal developing neurons RAGE
CC colocalizes with amphoterin which is a matrix associated polypeptide.
CC The expression of both RAGE and amphoterin decreases after birth, but
CC both have increased expression in tumours. RAGE polypeptides
CC AAY52132-Y52135 are used in the invention in a method for inhibiting
CC tumour invasion and metastasis. The method involves inhibiting tumour
CC invasion and metastasis via administration of a therapeutically effective
CC amount of the pharmaceutical composition containing a RAGE polypeptide.
CC The invention also relates to a method for evaluating the ability of an
CC agent to inhibit tumour invasion in a local cellular environment. RAGE
CC can be administered to a patient in a pharmaceutically acceptable
CC carrier.
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 30; DB 21; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.8e-25;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||||||||
Db 1 aqnitarigeplvlkckgapkkppqrlewk 30

RESULT 3
AAW44200
ID AAW44200 standard; Protein; 318 AA.
XX
AC AAW44200;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human mature receptor to an advanced glycosylation end product.
XX
KW Human; soluble receptor; advanced glycosylation end product; RAGE;
KW AGE; antibody; vascular permeability; diabetes mellitus.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 66
FT /note= "encoded by CCT"
XX
PN WO9739125-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01834.
XX
PR 16-APR-1996; 96US-0633148.
XX
PA (SCHD) SCHERING PATENTE AG.
XX
PI Hollander DA, Morser MJ, Nagashima M;
XX
DR WPI; 1997-558580/51.
DR N-PSDB; AAV12395.
XX
PT Anti-advanced glycosylation end product polypeptide antibody -
PT prevents receptor binding and therefore reduces vascular
PT permeability, useful to treat diabetes mellitus
XX
PS Claim 2; Page 42-43; 90pp; English.

XX
CC The present sequence represents a mature human receptor to an advanced
CC glycosylation end product (RAGE) polypeptide. The present invention
CC describes an isolated antibody (Ab), specifically immunoreactive with
CC RAGE. Advanced glycosylation end products (AGE) of proteins are
CC non-enzymatically glycosylated proteins, which accumulate in vascular
CC tissue in ageing, and at an accelerated rate in individuals with
CC diabetes. The Ab, which prevents the interaction between an AGE and it's
CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
CC associated amyloidosis or atherosclerosis. The Ab can also be used for
CC the isolation and purification of human RAGE polypeptide.
XX
SQ Sequence 318 AA;

Query Match 100.0%; Score 30; DB 18; Length 318;
Best Local Similarity 100.0%; Pred. No. 4e-24;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||||||||
Db 1 aqnitarigeplvlkckgapkkppqrlewk 30

RESULT 4
AAW33754
ID AAW33754 standard; Protein; 318 AA.
XX
AC AAW33754;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human RAGE polypeptide (318 amino acid residues).
XX
KW Advanced glycosylation end-product receptor; RAGE; screening; AGE;
KW vascular permeability; diabetes mellitus; treatment; atherosclerosis;
KW Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO9739121-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01832.
XX
PR 16-APR-1996; 96US-0633147.
XX
PA (SCHD) SCHERING AG.
XX
PI Morser MJ, Nagashima M;
XX
DR WPI; 1997-526458/48.
DR N-PSDB; AAV06518.
XX
PT New soluble advanced glycosylation end-product receptor polypeptide
PT - used for reducing vascular permeability, complications of diabetes
PT etc., also for purification and to screen for modulators
XX
PS Claim 3; Fig 1B; 91pp; English.
XX
CC This is a human advanced glycosylation end-product receptor (RAGE)
CC polypeptide (318 amino acid residues). The RAGE polypeptides and its
CC active fragments or their mimetics, inhibit interaction between advanced
CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They
CC are used to treat diseases associated with AGE/RAGE interaction, such as
CC increased vascular permeability, diabetes mellitus (particularly
CC complications such as micro- or macro- vasculopathy or occlusive vascular
CC disorders such as neuropathy, nephropathy, retinopathy or
CC atherosclerosis) or haemodialysis-associated amyloidosis, also activation
CC of microglial cells by beta-amyloid peptides in Alzheimer's disease or
CC age-related disorders such as oxidative stress. These RAGE polypeptides
CC are also used, when immobilised, to purify AGE from a protein mixture and
CC to screen for compounds that are agonists and antagonists of AGE/RAGE
CC interaction. They can also be used diagnostically to detect abnormal
CC levels of AGE. Antibodies against RAGE polypeptides are useful as
CC immunoassay reagents for measurement of RAGE levels, and as inhibitors of
CC interaction between AGE and RAGE or other receptors and for purification
CC and quantification of RAGE polypeptides. The encoding nucleic acids are
CC used to express recombinant RAGE and as probes for isolating related
CC genes.
XX
SQ Sequence 318 AA;

Query Match 100.0%; Score 30; DB 18; Length 318;
Best Local Similarity 100.0%; Pred. No. 4e-24;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||||||||
Db 1 aqnitarigeplvlkckgapkkppqrlewk 30

RESULT 5
AAY52130
ID AAY52130 standard; protein; 332 AA.
XX
AC AAY52130;
XX
DT 28-JAN-2000 (first entry)
XX
DE Human Receptor to AGE (RAGE) amino acid sequence.
XX
KW Soluble receptor for advanced glycation endproducts; RAGE; tumour;
KW invasion; metastasis; amphoterin; neuron; inhibit; therapy.
XX
OS Homo sapiens.
XX
PN WO9954485-A1.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US08427.
XX
PR 17-APR-1998; 98US-0062365.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Schmidt AM, Stern D;
XX
DR WPI; 2000-013260/01.
XX
PT Inhibiting tumour invasion or spreading by administration of soluble
PT receptor for advanced glycation endproducts -
XX
PS Disclosure; Page 10-11; 88pp; English.
XX
CC This is the amino acid sequence of the human soluble Receptor for
CC Advanced Glycation Endproducts (RAGE). RAGE interacts with a range of
CC physiologically and pathophysiologically relevant ligands when
CC considering tumour invasion. In normal developing neurons RAGE
CC colocalizes with amphoterin which is a matrix associated polypeptide.
CC The expression of both RAGE and amphoterin decreases after birth, but
CC both have increased expression in tumours. RAGE polypeptides
CC AAY52132-Y52135 are used in the invention in a method for inhibiting
CC tumour invasion and metastasis. The method involves inhibiting tumour
CC invasion and metastasis via administration of a therapeutically effective
CC amount of the pharmaceutical composition containing a RAGE polypeptide.
CC The invention also relates to a method for evaluating the ability of an
CC agent to inhibit tumour invasion in a local cellular environment. RAGE
CC can be administered to a patient in a pharmaceutically acceptable
CC carrier.
XX

```
SQ   Sequence   332 AA;


  Query Match             100.0%;  Score 30;  DB 21;  Length 332;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-24;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
          ||||||||||||||||||||||||||||||
Db      1 aqnitarigeplvlkckgapkkppqrlewk 30


RESULT   6
AAW44199
ID   AAW44199 standard; Protein; 340 AA.
XX
AC   AAW44199;
XX
DT   14-MAY-1998  (first entry)
XX
DE   Human soluble receptor to an advanced glycosylation end product.
XX
KW   Human; soluble receptor; advanced glycosylation end product; RAGE;
KW   AGE; antibody; vascular permeability; diabetes mellitus.
XX
OS   Homo sapiens.
XX
PN   WO9739125-A1.
XX
PD   23-OCT-1997.
XX
PF   11-APR-1997;   97WO-EP01834.
XX
PR   16-APR-1996;   96US-0633148.
XX
PA   (SCHD ) SCHERING PATENTE AG.
XX
PI   Hollander DA,  Morser MJ,  Nagashima M;
XX
DR   WPI; 1997-558580/51.
DR   N-PSDB; AAV12394.
XX
PT   Anti-advanced glycosylation end product polypeptide antibody -
PT   prevents receptor binding and therefore reduces vascular
PT   permeability, useful to treat diabetes mellitus
XX
PS   Claim 2; Page 40-41; 90pp; English.
XX
CC   The present sequence represents a soluble human receptor to an advanced
CC   glycosylation end product (RAGE) polypeptide. The present invention
CC   describes an isolated antibody (Ab), specifically immunoreactive with
CC   RAGE. Advanced glycosylation end products (AGE) of proteins are
CC   non-enzymatically glycosylated proteins, which accumulate in vascular
CC   tissue in ageing, and at an accelerated rate in individuals with
CC   diabetes. The Ab, which prevents the interaction between an AGE and it's
CC   receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC   treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC   vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
CC   associated amyloidosis or atherosclerosis. The Ab can also be used for
CC   the isolation and purification of human RAGE polypeptide.
XX
SQ   Sequence   340 AA;


  Query Match             100.0%;  Score 30;  DB 18;  Length 340;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-24;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
          ||||||||||||||||||||||||||||||
Db     23 aqnitarigeplvlkckgapkkppqrlewk 52


RESULT   7
AAW33753
ID   AAW33753 standard; Protein; 340 AA.
XX
AC   AAW33753;
XX
DT   08-MAY-1998  (first entry)
XX
DE   Human RAGE polypeptide (340 amino acid residues).
XX
KW   Advanced glycosylation end-product receptor; RAGE; screening; AGE;
KW   vascular permeability; diabetes mellitus; treatment; atherosclerosis;
KW   Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO9739121-A1.
XX
PD   23-OCT-1997.
XX
PF   11-APR-1997;   97WO-EP01832.
XX
PR   16-APR-1996;   96US-0633147.
XX
PA   (SCHD ) SCHERING AG.
XX
PI   Morser MJ,  Nagashima M;
XX
DR   WPI; 1997-526458/48.
DR   N-PSDB; AAV06517.
XX
PT   New soluble advanced glycosylation end-product receptor polypeptide
PT   - used for reducing vascular permeability, complications of diabetes
PT   etc., also for purification and to screen for modulators
XX
PS   Claim 3; Fig 1A; 91pp; English.
XX
CC   This is a human advanced glycosylation end-product receptor (RAGE)
CC   polypeptide (340 amino acid residues). The RAGE polypeptides and its
CC   active fragments or their mimetics, inhibit interaction between advanced
CC   glycosylation end-products (AGE) and a receptor (specifically RAGE). They
CC   are used to treat diseases associated with AGE/RAGE interaction, such as
CC   increased vascular permeability, diabetes mellitus (particularly
CC   complications such as micro- or macro- vasculopathy or occlusive vascular
CC   disorders such as neuropathy, nephropathy, retinopathy or
CC   atherosclerosis) or haemodialysis-associated amyloidosis, also activation
CC   of microglial cells by beta-amyloid peptides in Alzheimer's disease or
CC   age-related disorders such as oxidative stress. These RAGE polypeptides
CC   are also used, when immobilised, to purify AGE from a protein mixture and
CC   to screen for compounds that are agonists and antagonists of AGE/RAGE
CC   interaction. They can also be used diagnostically to detect abnormal
CC   levels of AGE. Antibodies against RAGE polypeptides are useful as
CC   immunoassay reagents for measurement of RAGE levels, and as inhibitors of
CC   interaction between AGE and RAGE or other receptors and for purification
CC   and quantification of RAGE polypeptides. The encoding nucleic acids are
CC   used to express recombinant RAGE and as probes for isolating related
CC   genes.
XX
SQ   Sequence   340 AA;


  Query Match             100.0%;  Score 30;  DB 18;  Length 340;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-24;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
          ||||||||||||||||||||||||||||||
Db     23 aqnitarigeplvlkckgapkkppqrlewk 52


RESULT   8
AAB81925
ID   AAB81925 standard; protein; 404 AA.
```

XX
AC AAB81925;
XX
DT 15-JUN-2001 (first entry)
XX
DE Extracorporeal circulation material receptor protein.
XX
KW Extracorporeal circulation; carbonyl stress product; receptor;
KW diabetes; vascular lesion; excretory dysfunction.
XX
OS Unidentified.
XX
PN WO200118060-A1.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-JP06172.
XX
PR 08-SEP-1999; 99JP-0254463.
XX
PA (TORA) TORAY IND INC.
XX
PI Shimizu S, Kubota M, Akiyama H, Usui M;
XX
DR WPI; 2001-290314/30.
XX
PT Material for extracorporeal circulation, applicable in selective
PT elimination of diabetic complication factors such as carbonyl stress
PT products caused by abnormally promoted carbonyl stress from excretory
PT dysfunction in vascular lesions -
XX
PS Claim 1; Page 31-32; 36pp; Japanese.
XX
CC The present invention describes a material for extracorporeal circulation
CC which is made from a water-insoluble carrier immobilized with a protein
CC having the sequence shown here. The materials of the invention, including
CC adsorbents, are for extracorporeal circulation, which are applicable in
CC the selective elimination of diabetic complication factors from a body
CC fluid, and are therefore useful in treating vascular lesions like
CC arteriosclerosis due to carbonyl stress products caused by abnormally
CC promoted carbonyl stress from excretory dysfunction.
XX
SQ Sequence 404 AA;

Query Match 100.0%; Score 30; DB 22; Length 404;
Best Local Similarity 100.0%; Pred. No. 4.9e-24;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||||||||
Db 23 aqnitarigeplvlkckgapkkppqrlewk 52

RESULT 9
AAW44214
ID AAW44214 standard; peptide; 16 AA.
XX
AC AAW44214;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human soluble RAGE immunologically active fragment SEQ ID NO:18.
XX
KW Human; soluble receptor; advanced glycosylation end product; RAGE;
KW AGE; antibody; vascular permeability; immunologically active fragment;
KW diabetes mellitus.
XX
OS Homo sapiens.
XX
PN WO9739125-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01834.
XX
PR 16-APR-1996; 96US-0633148.
XX
PA (SCHD) SCHERING PATENTE AG.
XX
PI Hollander DA, Morser MJ, Nagashima M;
XX
DR WPI; 1997-558580/51.
XX
PT Anti-advanced glycosylation end product polypeptide antibody -
PT prevents receptor binding and therefore reduces vascular
PT permeability, useful to treat diabetes mellitus
XX
PS Claim 2; Page 49; 90pp; English.
XX
CC The present sequence represents an immunologically active fragment
CC of a soluble human receptor to an advanced glycosylation end
CC product (RAGE) polypeptide. The present invention describes
CC an isolated antibody (Ab), specifically immunoreactive with
CC RAGE. Advanced glycosylation end products (AGE) of proteins are
CC non-enzymatically glycosylated proteins, which accumulate in vascular
CC tissue in ageing, and at an accelerated rate in individuals with
CC diabetes. The Ab, which prevents the interaction between an AGE and it's
CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
CC associated amyloidosis or atherosclerosis. The Ab can also be used for
CC the isolation and purification of human RAGE polypeptide.
XX
SQ Sequence 16 AA;

Query Match 53.3%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKC 16
||||||||||||||||
Db 1 aqnitarigeplvlkc 16

RESULT 10
AAW33768
ID AAW33768 standard; peptide; 16 AA.
XX
AC AAW33768;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human RAGE polypeptide fragment 14.
XX
KW Advanced glycosylation end-product receptor; RAGE; screening; AGE;
KW vascular permeability; diabetes mellitus; treatment; atherosclerosis;
KW Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO9739121-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01832.
XX
PR 16-APR-1996; 96US-0633147.
XX
PA (SCHD) SCHERING AG.
XX
PI Morser MJ, Nagashima M;
XX
DR WPI; 1997-526458/48.
XX

PT New soluble advanced glycosylation end-product receptor polypeptide
PT - used for reducing vascular permeability, complications of diabetes
PT etc., also for purification and to screen for modulators
XX
PS Disclosure; Page 9; 91pp; English.
XX
CC This is a peptide fragment of a human advanced glycosylation end-product
CC receptor (RAGE) polypeptide. The RAGE polypeptides and its active
CC fragments or their mimetics can inhibit interaction between advanced
CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They
CC are used to treat diseases associated with AGE/RAGE interaction, such as
CC increased vascular permeability, diabetes mellitus (particularly
CC complications such as micro- or macro- vasculopathy or occlusive vascular
CC disorders such as neuropathy, nephropathy, atherosclerosis or
CC retinopathy) or haemodialysis-associated amyloidosis, also activation
CC of microglial cells by beta-amyloid peptides in Alzheimer's disease or
CC age-related disorders such as oxidative stress. These RAGE polypeptides
CC are also used, when immobilised, to purify AGE from a protein mixture and
CC to screen for compounds that are agonists and antagonists of AGE/RAGE
CC interaction. They can also be used diagnostically to detect abnormal
CC levels of AGE. Antibodies against RAGE polypeptides are useful as
CC immunoassay reagents for measurement of RAGE levels, and as inhibitors of
CC interaction between AGE and RAGE or other receptors and for purification
CC and quantification of RAGE polypeptides. The encoding nucleic acids are
CC used to express recombinant RAGE and as probes for isolating related
CC genes.
XX
SQ Sequence 16 AA;

Query Match 53.3%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKC 16
||||||||||||||||
Db 1 aqnitarigeplvlkc 16

RESULT 11
AAW44208
ID AAW44208 standard; peptide; 15 AA.
XX
AC AAW44208;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human soluble RAGE immunologically active fragment SEQ ID NO:12.
XX
KW Human; soluble receptor; advanced glycosylation end product; RAGE;
KW AGE; antibody; vascular permeability; immunologically active fragment;
KW diabetes mellitus.
XX
OS Homo sapiens.
XX
PN WO9739125-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01834.
XX
PR 16-APR-1996; 96US-0633148.
XX
PA (SCHD) SCHERING PATENTE AG.
XX
PI Hollander DA, Morser MJ, Nagashima M;
XX
DR WPI; 1997-558580/51.
XX
PT Anti-advanced glycosylation end product polypeptide antibody -
PT prevents receptor binding and therefore reduces vascular
PT permeability, useful to treat diabetes mellitus
XX
PS Claim 2; Page 46; 90pp; English.
XX
CC The present sequence represents an immunologically active fragment
CC of a soluble human receptor to an advanced glycosylation end
CC product (RAGE) polypeptide. The present invention describes
CC an isolated antibody (Ab), specifically immunoreactive with
CC RAGE. Advanced glycosylation end products (AGE) of proteins are
CC non-enzymatically glycosylated proteins, which accumulate in vascular
CC tissue in ageing, and at an accelerated rate in individuals with
CC diabetes. The Ab, which prevents the interaction between an AGE and it's
CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
CC associated amyloidosis or atherosclerosis. The Ab can also be used for
CC the isolation and purification of human RAGE polypeptide.
XX
SQ Sequence 15 AA;

Query Match 50.0%; Score 15; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLK 15
|||||||||||||||
Db 1 aqnitarigeplvlk 15

RESULT 12
AAW33762
ID AAW33762 standard; peptide; 15 AA.
XX
AC AAW33762;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human RAGE polypeptide fragment 8.
XX
KW Advanced glycosylation end-product receptor; RAGE; screening; AGE;
KW vascular permeability; diabetes mellitus; treatment; atherosclerosis;
KW Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO9739121-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01832.
XX
PR 16-APR-1996; 96US-0633147.
XX
PA (SCHD) SCHERING AG.
XX
PI Morser MJ, Nagashima M;
XX
DR WPI; 1997-526458/48.
XX
PT New soluble advanced glycosylation end-product receptor polypeptide
PT - used for reducing vascular permeability, complications of diabetes
PT etc., also for purification and to screen for modulators
XX
PS Claim 6; Page 55; 91pp; English.
XX
CC This is a peptide fragment of a human advanced glycosylation end-product
CC receptor (RAGE) polypeptide. The RAGE polypeptides and its active
CC fragments or their mimetics can inhibit interaction between advanced
CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They
CC are used to treat diseases associated with AGE/RAGE interaction, such as
CC increased vascular permeability, diabetes mellitus (particularly
CC complications such as micro- or macro- vasculopathy or occlusive vascular
CC disorders such as neuropathy, nephropathy, atherosclerosis or
CC retinopathy) or haemodialysis-associated amyloidosis, also activation

CC of microglial cells by beta-amyloid peptides in Alzheimer's disease or
CC age-related disorders such as oxidative stress. These RAGE polypeptides
CC are also used, when immobilised, to purify AGE from a protein mixture and
CC to screen for compounds that are agonists and antagonists of AGE/RAGE
CC interaction. They can also be used diagnostically to detect abnormal
CC levels of AGE. Antibodies against RAGE polypeptides are useful as
CC immunoassay reagents for measurement of RAGE levels, and as inhibitors of
CC interaction between AGE and RAGE or other receptors and for purification
CC and quantification of RAGE polypeptides. The encoding nucleic acids are
CC used to express recombinant RAGE and as probes for isolating related
CC genes.
XX
SQ Sequence 15 AA;

Query Match 50.0%; Score 15; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLK 15
||||||||||||||||
Db 1 aqnitarigeplvlk 15

RESULT 13
AAY09350
ID AAY09350 standard; peptide; 30 AA.
XX
AC AAY09350;
XX
DT 09-JUL-1999 (first entry)
XX
DE Mouse RAGE V-domain peptide SEQ ID NO:2.
XX
KW RAGE; V-domain; receptor for advanced glycation endproduct;
KW ligand binding site; amyloid-beta; Alzheimer's disease; Down's syndrome;
KW senility; renal failure; hyperlipidaemic atherosclerosis; dementia;
KW neuronal cytotoxicity; head trauma; amyotrophic lateral sclerosis;
KW multiple sclerosis; amyloidosis; autoimmune disease; inflammation;
KW tumour; cancer; male impotence; wound healing; periodontal disease;
KW neuropathy; retinopathy; nephropathy; neuronal degeneration.
XX
OS Mus sp.
XX
PN WO9918987-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21346.
XX
PR 09-OCT-1997; 97US-0948131.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Lamster I, Schmidt AM, Stern D, Yan SD;
XX
DR WPI; 1999-277439/23.
XX
PT New peptides based on an advanced glycation end product receptor are
PT useful for treating Alzheimer's disease and Down's syndrome
XX
PS Claim 3; Page 78; 101pp; English.
XX
CC The present invention describes novel isolated peptides (I) having an
CC amino acid sequence corresponding to an amino acid sequence of a
CC V-domain of a receptor for an advanced glycation end product (RAGE).
CC Also described are methods for: (1) inhibiting an amyloid-beta peptide
CC (ABP) interaction with a receptor for RAGE when the receptor is on the
CC surface of a cell; (2) inhibiting degeneration of a neuronal cell; (3)
CC inhibiting formation of an ABP fibril on a cell; (4) inhibiting
CC extracellular assembly of an ABP into a fibril; (5) inhibiting
CC aggregation of ABP on the surface of a cell; (6) inhibiting infiltration
CC of a microglial cell into senile plaques; (7) inhibiting activation of a
CC microglial cell by an ABP; (8) treating a subject with a condition
CC associated with an interaction of an ABP with a receptor for RAGE on a
CC cell; (9) evaluating the ability of an agent to inhibit binding of an
CC ABP with a V-domain of a receptor for RAGE on the surface of a cell; (10)
CC inhibiting activation of a NF-kappaB gene in a cell; (11) inhibiting
CC periodontal disease in a subject; (12) inhibiting an RAGE's interaction
CC with a receptor for RAGE when the receptor is on the surface of a cell;
CC and (13) treating a subject with a condition associated with an
CC interaction of an RAGE with a receptor for RAGE on a cell. The methods
CC can be used for treating conditions associated with an interaction of an
CC ABP or an RAGE with a receptor for RAGE, e.g. diabetes, Alzheimer's
CC disease, senility, renal failure, hyperlipidaemic atherosclerosis,
CC neuronal cytotoxicity, Down's syndrome, dementia associated with head
CC trauma, amyotrophic lateral sclerosis, multiple sclerosis, amyloidosis,
CC an autoimmune disease, inflammation, a tumour, cancer, male impotence,
CC wound healing, periodontal disease, neuropathy, retinopathy, nephropathy
CC or neuronal degeneration.
XX
SQ Sequence 30 AA;

Query Match 43.3%; Score 13; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QNITARIGEPLVL 14
|||||||||||||
Db 2 qnitarigeplvl 14

RESULT 14
AAY09351
ID AAY09351 standard; peptide; 30 AA.
XX
AC AAY09351;
XX
DT 09-JUL-1999 (first entry)
XX
DE Rat RAGE V-domain peptide SEQ ID NO:3.
XX
KW RAGE; V-domain; receptor for advanced glycation endproduct;
KW ligand binding site; amyloid-beta; Alzheimer's disease; Down's syndrome;
KW senility; renal failure; hyperlipidaemic atherosclerosis; dementia;
KW neuronal cytotoxicity; head trauma; amyotrophic lateral sclerosis;
KW multiple sclerosis; amyloidosis; autoimmune disease; inflammation;
KW tumour; cancer; male impotence; wound healing; periodontal disease;
KW neuropathy; retinopathy; nephropathy; neuronal degeneration.
XX
OS Rattus sp.
XX
PN WO9918987-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21346.
XX
PR 09-OCT-1997; 97US-0948131.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Lamster I, Schmidt AM, Stern D, Yan SD;
XX
DR WPI; 1999-277439/23.
XX
PT New peptides based on an advanced glycation end product receptor are
PT useful for treating Alzheimer's disease and Down's syndrome
XX
PS Claim 4; Page 78; 101pp; English.
XX
CC The present invention describes novel isolated peptides (I) having an
CC amino acid sequence corresponding to an amino acid sequence of a
CC V-domain of a receptor for an advanced glycation end product (RAGE).
CC Also described are methods for: (1) inhibiting an amyloid-beta peptide

CC (ABP) interaction with a receptor for RAGE when the receptor is on the
CC surface of a cell; (2) inhibiting degeneration of a neuronal cell; (3)
CC inhibiting formation of an ABP fibril on a cell; (4) inhibiting
CC extracellular assembly of an ABP into a fibril; (5) inhibiting
CC aggregation of ABP on the surface of a cell; (6) inhibiting infiltration
CC of a microglial cell into senile plaques; (7) inhibiting activation of a
CC microglial cell by an ABP; (8) treating a subject with a condition
CC associated with an interaction of an ABP with a receptor for RAGE on a
CC cell; (9) evaluating the ability of an agent to inhibit binding of an
CC ABP with a V-domain of a receptor for RAGE on the surface of a cell; (10)
CC inhibiting activation of a NF-kappaB gene in a cell; (11) inhibiting
CC periodontal disease in a subject; (12) inhibiting an RAGE's interaction
CC with a receptor for RAGE when the receptor is on the surface of a cell;
CC and (13) treating a subject with a condition associated with an
CC interaction of an RAGE with a receptor for RAGE on a cell. The methods
CC can be used for treating conditions associated with an interaction of an
CC ABP or an RAGE with a receptor for RAGE, e.g. diabetes, Alzheimer's
CC disease, senility, renal failure, hyperlipidaemic atherosclerosis,
CC neuronal cytotoxicity, Down's syndrome, dementia associated with head
CC trauma, amyotrophic lateral sclerosis, multiple sclerosis, amyloidosis,
CC an autoimmune disease, inflammation, a tumour, cancer, male impotence,
CC wound healing, periodontal disease, neuropathy, retinopathy, nephropathy
CC or neuronal degeneration.
XX
SQ Sequence 30 AA;

Query Match 36.7%; Score 11; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QNITARIGEPL 12
||||||||||
Db 2 qnitarigepl 12

RESULT 15
AAW44209
ID AAW44209 standard; peptide; 10 AA.
XX
AC AAW44209;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human soluble RAGE immunologically active fragment SEQ ID NO:13.
XX
KW Human; soluble receptor; advanced glycosylation end product; RAGE;
KW AGE; antibody; vascular permeability; immunologically active fragment;
KW diabetes mellitus.
XX
OS Homo sapiens.
XX
PN WO9739125-A1.
XX
PD 23-OCT-1997.
XX
PF 11-APR-1997; 97WO-EP01834.
XX
PR 16-APR-1996; 96US-0633148.
XX
PA (SCHD) SCHERING PATENTE AG.
XX
PI Hollander DA, Morser MJ, Nagashima M;
XX
DR WPI; 1997-558580/51.
XX
PT Anti-advanced glycosylation end product polypeptide antibody
PT prevents receptor binding and therefore reduces vascular
PT permeability, useful to treat diabetes mellitus
XX
PS Claim 2; Page 47; 90pp; English.
XX
CC The present sequence represents an immunologically active fragment
CC of a soluble human receptor to an advanced glycosylation end
CC product (RAGE) polypeptide. The present invention describes
CC an isolated antibody (Ab), specifically immunoreactive with
CC RAGE. Advanced glycosylation end products (AGE) of proteins are
CC non-enzymatically glycosylated proteins, which accumulate in vascular
CC tissue in ageing, and at an accelerated rate in individuals with
CC diabetes. The Ab, which prevents the interaction between an AGE and it's
CC receptor (RAGE), reduces vascular permeability. The Ab can be used to
CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive
CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis
CC associated amyloidosis or atherosclerosis. The Ab can also be used for
CC the isolation and purification of human RAGE polypeptide.
XX
SQ Sequence 10 AA;

Query Match 33.3%; Score 10; DB 18; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 CKGAPKKPPQ 25
||||||||||
Db 1 ckgapkkppq 10

Search completed: July 31, 2002, 15:05:35
Job time: 110 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:03:45 ; Search time 14.58 Seconds
(without alignments)
197.715 Million cell updates/sec

Title: US-08-948-131-1
Perfect score: 30
Sequence: 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	404	1	I61596	advanced glycosyla
2	13	43.3	402	2	T09062	probable advanced
3	10	33.3	416	1	A42879	advanced glycosyla
4	7	23.3	223	2	H64065	mutator mutH - Hae
5	7	23.3	330	2	E81998	pseudouridylate sy
6	7	23.3	330	2	H81225	ribosomal large ch
7	7	23.3	537	1	JQ1619	cell fusion glycop
8	7	23.3	594	2	AB3343	single-stranded-DN
9	7	23.3	595	2	B86212	protein F24B9.20 [
10	6	20.0	96	2	T07822	cystein proteinase
11	6	20.0	110	2	S64538	hypothetical prote
12	6	20.0	155	2	T46154	hypothetical prote
13	6	20.0	159	2	D97739	30S ribosomal prot
14	6	20.0	201	2	AI3484	low molecular weig
15	6	20.0	204	2	A96024	probable acetyltra
16	6	20.0	209	2	E70536	hypothetical prote
17	6	20.0	230	2	A75260	hypothetical prote
18	6	20.0	234	1	QQBE43	membrane antigen g
19	6	20.0	238	2	H64181	ABC-type transport
20	6	20.0	242	2	AC2534	hypothetical prote
21	6	20.0	257	2	D70552	hypothetical prote
22	6	20.0	259	2	G86962	probable membrane
23	6	20.0	268	2	AI2237	septum site-determ
24	6	20.0	269	2	AF3494	molybdopterin bios
25	6	20.0	271	2	A70415	molybdopterin bios
26	6	20.0	282	2	T07303	cell division inhi
27	6	20.0	298	2	D87450	transcription regu
28	6	20.0	298	2	T51746	RNA helicase RH17
29	6	20.0	314	2	AH1912	hypothetical prote
30	6	20.0	326	2	H83187	hypothetical prote
31	6	20.0	333	2	AH1361	coat protein [Bact
32	6	20.0	334	2	D83164	conserved hypothet
33	6	20.0	336	2	B98209	hypothetical prote
34	6	20.0	336	2	AG3077	hypothetical prote
35	6	20.0	359	2	T35179	oxidoreductase alp
36	6	20.0	377	2	D97102	uncharacterized pr
37	6	20.0	387	2	C96670	hypothetical prote
38	6	20.0	389	2	S33667	probable integrase
39	6	20.0	429	2	T23984	hypothetical prote
40	6	20.0	435	2	AD2583	two component sens
41	6	20.0	468	2	H69133	argininosuccinate
42	6	20.0	471	2	S30585	hypothetical prote
43	6	20.0	474	2	B82227	exodeoxyribonuclea
44	6	20.0	479	2	C86275	hypothetical prote
45	6	20.0	481	2	S43687	serotonin receptor

ALIGNMENTS

RESULT 1
I61596
advanced glycosylation end-products receptor precursor - human
N;Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprot
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C;Accession: I61596; B42879; S27968
R;Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko,
Genomics 23, 408-419, 1994
A;Title: Three genes in the human MHC class III region near the junction with the cla
nterpart of mouse mammary tumor gene int-3.
A;Reference number: A55562; MUID:95137587
A;Accession: I61596
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A;Molecule type: DNA
A;Residues: 1-404 <RES>
A;Cross-references: GB:D28769; NID:g561657; PIDN:BAA05958.1; PID:g561659
R;Neeper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.;
J. Biol. Chem. 267, 14998-15004, 1992
A;Title: Cloning and expression of a cell surface receptor for advanced glycosylation
A;Reference number: A42879; MUID:92340547
A;Accession: B42879
A;Molecule type: mRNA
A;Residues: 'G',2-99,'R',101-404 <NEE>
A;Cross-references: EMBL:M91211; NID:g190845; PIDN:AAA03574.1; PID:g190846
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBIP:109438)
C;Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly
cellular function, thus contributing to tissue lesions in diabetes.
C;Comment: This receptor appears also to mediate the effects of amyloid beta peptide
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C;Genetics:
A;Gene: GDB:AGER
A;Cross-references: GDB:306354; OMIM:600214
A;Map position: 6p21.3-6p21.3
A;Introns: 18/1; 53/3; 119/1; 140/3; 170/1; 231/1; 274/3; 322/1; 331/1; 373/2
C;Function:
A;Description: neuronal receptor for amphoterin, a DNA-binding protein involved in ne
C;Superfamily: advanced glycosylation end products receptor; immunoglobulin homology
C;Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-404/Product: advanced glycosylation end products receptor #status predicted <MAT
F;23-344/Domain: extracellular #status predicted <EXT>
F;31-101/Domain: immunoglobulin homology <IM1>
F;137-210/Domain: immunoglobulin homology <IM2>
F;252-303/Domain: immunoglobulin homology <IM3>
F;345-362/Domain: transmembrane #status predicted <TMM>
F;363-404/Domain: intracellular #status predicted <INT>
F;25,81/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;38-99,144-208,259-301/Disulfide bonds: #status predicted

```
  Query Match             100.0%;  Score 30;  DB 1;  Length 404;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-23;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
          ||||||||||||||||||||||||||||||
Db     23 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 52


RESULT   2
T09062
probable advanced glycosylation end-products receptor precursor - mouse
N;Alternate names: RAGE
C;Species: Mus musculus (house mouse)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000
C;Accession: T09062
R;Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Sc
submitted to the EMBL Data Library, October 1997
A;Description: Sequence of the mouse major histocompatibility locus class III region.
A;Reference number: Z16543
A;Accession: T09062
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-402 <ROW>
A;Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564950
C;Genetics:
A;Gene: RAGE
A;Map position: 17
A;Introns: 18/1; 53/3; 118/1; 139/3; 169/1; 229/1; 272/3; 320/1; 329/1; 371/2
C;Superfamily: advanced glycosylation end products receptor; immunoglobulin homology
C;Keywords: receptor; transmembrane protein
F;31-100/Domain: immunoglobulin homology <IMM>


  Query Match             43.3%;  Score 13;  DB 2;  Length 402;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 QNITARIGEPLVL 14
          |||||||||||||
Db     24 QNITARIGEPLVL 36


RESULT   3
A42879
advanced glycosylation end-products receptor precursor - bovine
N;Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprotein
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 07-Feb-1997 #text_change 16-Jul-1999
C;Accession: A42879; A42878; S27949
R;Neeper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; St
J. Biol. Chem. 267, 14998-15004, 1992
A;Title: Cloning and expression of a cell surface receptor for advanced glycosylation en
A;Reference number: A42879; MUID:92340547
A;Accession: A42879
A;Molecule type: mRNA
A;Residues: 1-416 <NEE>
A;Cross-references: GB:M91212; NID:g163650; PIDN:AAA03575.1; PID:g163651
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBIP:109436)
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Schmidt, A.M.; Vianna, M.; Gerlach, M.; Brett, J.; Ryan, J.; Kao, J.; Esposito, C.; He
J. Biol. Chem. 267, 14987-14997, 1992
A;Title: Isolation and characterization of two binding proteins for advanced glycosylati
A;Reference number: A42878; MUID:92340546
A;Accession: A42878
A;Molecule type: protein
A;Residues: 23-24,'X',26-37,'X',39-49,'XX',52-54 <SCH>
A;Experimental source: endothelial cells
A;Note: sequence extracted from NCBI backbone (NCBIP:109434)
C;Comment: Advanced glycosylation end products are heterogeneous nonenzymatically glycos
cellular function, thus contributing to tissue lesions in diabetes.
C;Comment: This receptor appears also to mediate the effects of amyloid beta peptide on
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C;Function:
A;Description: neuronal receptor for amphoterin, a DNA-binding protein involved in ne
C;Superfamily: advanced glycosylation end products receptor; immunoglobulin homology
C;Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-416/Product: advanced glycosylation end-products receptor RAGE #status predicted
F;23-354/Domain: extracellular #status predicted <EXT>
F;31-100/Domain: immunoglobulin homology <IM1>
F;136-209/Domain: immunoglobulin homology <IM2>
F;262-313/Domain: immunoglobulin homology <IM3>
F;355-372/Domain: transmembrane #status predicted <TMM>
F;373-416/Domain: intracellular #status predicted <INT>
F;25,80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;38-98,143-207,269-311/Disulfide bonds: #status predicted


  Query Match             33.3%;  Score 10;  DB 1;  Length 416;
  Best Local Similarity   100.0%;  Pred. No. 0.0084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     16 CKGAPKKPPQ 25
          ||||||||||
Db     38 CKGAPKKPPQ 47


RESULT   4
H64065
mutator mutH - Haemophilus influenzae (strain Rd KW20)
N;Alternate names: DNA mismatch protein
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 26-Aug-1999
C;Accession: H64065
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630
A;Accession: H64065
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <TIGR>
A;Cross-references: GB:U32723; GB:L42023; NID:g1573363; PIDN:AAC22062.1; PID:g1573374
C;Genetics:
A;Gene: mutH
C;Superfamily: mutator mutH


  Query Match             23.3%;  Score 7;  DB 2;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 6.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 ITARIGE 10
          |||||||
Db    165 ITARIGE 171


RESULT   5
E81998
pseudouridylate synthase (EC 4.2.1.70) C rRNA-specific NMA0070 [imported] - Neisseria
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: E81998
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491
A;Reference number: A81775; MUID:20222556
A;Accession: E81998
A;Status: preliminary
A;Molecule type: DNA
```

A;Residues: 1-330 <PAR>
A;Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83386.1; PID:g737884
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: rluC; NMA0070
C;Superfamily: conserved hypothetical protein HI0176
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 23.3%; Score 7; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 GEPLVLK 15
|||||||
Db 304 GEPLVLK 310

RESULT 6
H81225
ribosomal large chain pseudouridine synthase C NMB0198 [imported] - Neisseria meningitid
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: H81225
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: H81225
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-330 <TET>
A;Cross-references: GB:AE002377; GB:AE002098; NID:g7225416; PIDN:AAF40655.1; PID:g722541
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0198
C;Superfamily: conserved hypothetical protein HI0176

Query Match 23.3%; Score 7; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 GEPLVLK 15
|||||||
Db 304 GEPLVLK 310

RESULT 7
JQ1619
cell fusion glycoprotein precursor - pneumonia virus of mice
N;Alternate names: F protein
N;Contains: cell fusion glycoprotein F1; cell fusion glycoprotein F2
C;Species: pneumonia virus of mice
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 25-Oct-1996
C;Accession: JQ1619
R;Chambers, P.; Pringle, C.R.; Easton, A.J.
J. Gen. Virol. 73, 1717-1724, 1992
A;Title: Sequence analysis of the gene encoding the fusion glycoprotein of pneumonia vir
A;Reference number: JQ1619; MUID:92333256
A;Accession: JQ1619
A;Molecule type: mRNA
A;Residues: 1-537 <CHA>
A;Cross-references: GB:S40186
C;Genetics:
A;Gene: F
C;Superfamily: cell fusion glycoprotein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-101/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F;102-537/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F;491-514/Domain: transmembrane #status predicted <TMM>
F;463,488/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.3%; Score 7; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 GEPLVLK 15
|||||||
Db 434 GEPLVLK 440

RESULT 8
AB3343
single-stranded-DNA-specific exonuclease recj (EC 3.1.-.-) [imported] - Brucella meli
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AB3343
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3343
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51909.1; PID:g17982663; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0728
A;Map position: I
C;Superfamily: single-stranded-DNA-specific exonuclease RecJ

Query Match 23.3%; Score 7; DB 2; Length 594;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ARIGEPL 12
|||||||
Db 287 ARIGEPL 293

RESULT 9
B86212
protein F24B9.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: B86212
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallc
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: B86212
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-595 <STO>
A;Cross-references: GB:AE005172; NID:g8439898; PIDN:AAF75084.1; GSPDB:GN00141
C;Genetics:
A;Gene: F24B9.20
A;Map position: 1

```
  Query Match             23.3%;  Score 7;  DB 2;  Length 595;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 RIGEPLV 13
           |||||||
Db     550 RIGEPLV 556


RESULT  10
T07822
cystein proteinase inhibitor - cucumber
C;Species: Cucumis sativus (cucumber)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
C;Accession: T07822
R;Yamakawa, S.
submitted to the EMBL Data Library, May 1998
A;Description: Cystein proteinase inhibitor.
A;Reference number: Z16154
A;Accession: T07822
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-96 <YAM>
A;Cross-references: EMBL:AB014760; PIDN:BAA28867.1
A;Note: root-specific
C;Superfamily: cystatin; cystatin homology
C;Keywords: cysteine proteinase inhibitor


  Query Match             20.0%;  Score 6;  DB 2;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LVLKCK 17
           ||||||
Db      59 LVLKCK 64


RESULT  11
S64538
hypothetical protein YGR215w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G7821
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C;Accession: S64538
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64538
A;Molecule type: DNA
A;Residues: 1-110 <RIE>
A;Cross-references: EMBL:Z73000; NID:g1323386; PID:g1323387; GSPDB:GN00007; MIPS:YGR215w
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YGR215w
A;Map position: 7R
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YGR215w


  Query Match             20.0%;  Score 6;  DB 2;  Length 110;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKK 22
           ||||||
Db      88 KGAPKK 93


RESULT  12
T46154
hypothetical protein T4D2.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46154
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23025
A;Accession: T46154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-155 <NYA>
A;Cross-references: EMBL:AL132958
A;Experimental source: cultivar Columbia; BAC clone T4D2
C;Genetics:
A;Map position: 3
A;Introns: 118/3
A;Note: T4D2.10


  Query Match             20.0%;  Score 6;  DB 2;  Length 155;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 KCKGAP 20
           ||||||
Db     139 KCKGAP 144


RESULT  13
D97739
30S ribosomal protein S9 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C;Accession: D97739
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: D97739
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-159 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL02854.1; PID:g15619376; GSPDB:GN00173
C;Genetics:
A;Gene: rpsI


  Query Match             20.0%;  Score 6;  DB 2;  Length 159;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 EPLVLK 15
           ||||||
Db      18 EPLVLK 23


RESULT  14
AI3484
low molecular weight phosphotyrosine protein phosphatase (EC 3.1.3.48) [imported] - B
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AI3484
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AI3484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL53044.1; PID:g17983904; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1863
A;Map position: I
```

C;Keywords: phosphoric monoester hydrolase

Query Match 20.0%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITARIG 9
||||||
Db 140 ITARIG 145

RESULT 15
A96024
probable acetyltransferase protein (EC 2.3.1.-) [imported] - Sinorhizobium meliloti (str
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: A96024
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: A96024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-204 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49857.1; PID:g15141345; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20765
A;Genome: plasmid
C;Superfamily: Agrobacterium chloramphenicol acetyltransferase
C;Keywords: acyltransferase

Query Match 20.0%; Score 6; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TARIGE 10
||||||
Db 164 TARIGE 169

Search completed: July 31, 2002, 15:04:30
Job time: 45 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:04:35 ; Search time 10.35 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-08-948-131-1
Perfect score: 30
Sequence: 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	404	1	RAGE_HUMAN	Q15109 homo sapien
2	13	43.3	403	1	RAGE_MOUSE	Q62151 mus musculu
3	11	36.7	402	1	RAGE_RAT	Q63495 rattus norv
4	10	33.3	416	1	RAGE_BOVIN	Q28173 bos taurus
5	7	23.3	223	1	MUTH_HAEIN	P44688 haemophilus
6	7	23.3	537	1	VGLF_PVM	P35949 pneumonia v
7	6	20.0	110	1	YG4L_YEAST	P53305 saccharomyc
8	6	20.0	159	1	RS9_RICCN	Q92iv4 rickettsia
9	6	20.0	181	1	YCF4_GUITH	O78467 guillardia
10	6	20.0	210	1	YA79_HAEIN	P45023 haemophilus
11	6	20.0	234	1	VGP8_EBV	P03224 epstein-bar
12	6	20.0	259	1	E434_ADECT	P87568 canine aden
13	6	20.0	265	1	E434_ADECC	Q65962 canine aden
14	6	20.0	265	1	E434_ADECR	Q96690 canine aden
15	6	20.0	269	1	MIND_GUITH	O78436 guillardia
16	6	20.0	278	1	PROC_VIBAL	P52053 vibrio algi
17	6	20.0	282	1	MIND_CHLVU	P56346 chlorella v
18	6	20.0	401	1	ENO_THEAC	Q9hjt1 thermoplasm
19	6	20.0	429	1	YRM4_CAEEL	Q09415 caenorhabdi
20	6	20.0	468	1	ARLY_METTH	O26369 methanobact
21	6	20.0	481	1	5H2B_HUMAN	P41595 homo sapien
22	6	20.0	482	1	DUSA_HUMAN	Q9y6w6 homo sapien
23	6	20.0	504	1	5H2B_MOUSE	Q02152 mus musculu
24	6	20.0	513	1	SPT1_HUMAN	O43278 homo sapien
25	6	20.0	607	1	UVRC_PSEFL	P32966 pseudomonas
26	6	20.0	626	1	RN17_MOUSE	Q99mv7 mus musculu
27	6	20.0	635	1	SYT_RICPR	O05947 rickettsia
28	6	20.0	714	1	EFGL_MYCTU	O07170 mycobacteri
29	6	20.0	754	1	ASPH_BOVIN	Q28056 bos taurus
30	6	20.0	757	1	ASPH_HUMAN	Q12797 homo sapien
31	6	20.0	788	1	REC2_HAEIN	P44408 haemophilus
32	6	20.0	837	1	HFC1_HAEIN	P33397 haemophilus
33	6	20.0	837	1	HFC2_HAEIN	P45997 haemophilus
34	6	20.0	837	1	HFC3_HAEIN	P45998 haemophilus
35	6	20.0	850	1	DEXT_STRMU	Q54443 streptococc
36	6	20.0	879	1	YN65_YEAST	P42837 saccharomyc
37	6	20.0	1097	1	KF1D_RAT	O35787 rattus norv
38	6	20.0	1103	1	KF1C_HUMAN	O43896 homo sapien
39	6	20.0	1733	1	VNUA_PRVKA	P33485 pseudorabie
40	6	20.0	2142	1	BAT2_HUMAN	P48634 homo sapien
41	6	20.0	2210	1	RRPO_TACV	P20430 tacaribe vi
42	5	16.7	51	1	MLEV_MOUSE	P09542 mus musculu
43	5	16.7	87	1	VE4_HPV51	P26548 human papil
44	5	16.7	99	1	RS20_CHLPN	Q9z7f2 chlamydia p
45	5	16.7	111	1	HMGZ_DROME	Q06943 drosophila

ALIGNMENTS

RESULT 1
RAGE_HUMAN
ID RAGE_HUMAN STANDARD; PRT; 404 AA.
AC Q15109; Q15279; Q9Y3R3; Q9H2X7;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Advanced glycosylation end product-specific receptor precursor
DE (Receptor for advanced glycosylation end products).
GN AGER OR RAGE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=92340547; PubMed=1378843;
RA Neeper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
RA Elliston K., Stern D., Shaw A.;
RT "Cloning and expression of a cell surface receptor for advanced
RT glycosylation end products of proteins.";
RL J. Biol. Chem. 267:14998-15004(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=95137587; PubMed=7835890;
RA Sugaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,
RA Inoko H., Ikemura T.;
RT "Three genes in the human MHC class III region near the junction with
RT the class II: gene for receptor of advanced glycosylation end
RT products, PBX2 homeobox gene and a notch homolog, human counterpart
RT of mouse mammary tumor gene int-3.";
RL Genomics 23:408-419(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Spies T., Hood L.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.
RA Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,
RA Yamamoto H.;
RT "Molecular heterogeneity of the receptor for advanced glycation
RT endproducts.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,
RA Schuler A., Huber G.;
RT "cDNA cloning of a novel secreted isoform of the human Receptor for
RT Advanced Glycation End products (RAGE) and characterization of cells
RT co-expressing cell-surface scavenger receptors and Swedish mutant
RT amyloid precursor protein.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE OF 1-12 FROM N.A.

```
RA   Hudson B.I., Futers T.S.;
RT   "Novel polymorphisms in the receptor for advanced glycation
RT   end-products (RAGE) gene.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
CC       PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
CC       WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
CC       RATE IN DIABETES.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ISOFORM 1) AND
CC       SECRETED (ISOFORM 2).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/RAGESEC;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M91211; AAA03574.1; -.
DR   EMBL; D28769; BAA05958.1; -.
DR   EMBL; U89336; AAB47491.1; -.
DR   EMBL; AB036432; BAA89369.1; -.
DR   EMBL; AJ133822; CAB43108.1; -.
DR   EMBL; AF208289; AAG35728.1; -.
DR   MIM; 600214; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Alternative splicing; Polymorphism.
FT   SIGNAL       1    22       POTENTIAL.
FT   CHAIN       23   404       ADVANCED GLYCOSYLATION END PRODUCT-
FT                              SPECIFIC RECEPTOR.
FT   DOMAIN      23   342       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   343   363       POTENTIAL.
FT   DOMAIN     364   404       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      31   106       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN     137   214       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN     252   308       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID    38    99       POTENTIAL.
FT   DISULFID   144   208       POTENTIAL.
FT   DISULFID   259   301       POTENTIAL.
FT   CARBOHYD    25    25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    81    81       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DOMAIN     380   384       POLY-GLU.
FT   VARSPLIC    54    67       MISSING (IN ISOFORM 2).
FT   VARSPLIC   275   404       GVPLPLPPSPVLILPEIGPQDQGTYSCVATHSSHGPQESRA
FT                              VSISIIEPGEEGPTAGSVGGSGLGTLALALGILGGLGTAAL
FT                              LIGVILWQRRQRRGEERKAPENQEEEEERAELNQSEEPEAG
FT                              ESSTGGP -> VSDLERGAGRTRRGGANCRLCGRIRAGNSS
FT                              PGPGDPGRPGDSRPAHWGHLVAKAATPRRGEEGPRKPGGRG
FT                              GACRTESVGGT (IN ISOFORM 2).
FT   VARIANT    100   100       Q -> R.
FT                              /FTId=VAR_011338.
FT   CONFLICT     1     1       M -> G (IN REF. 1).
SQ   SEQUENCE   404 AA;  42802 MW;  0D584C436C30CCE7 CRC64;
```

```
Query Match          100.0%;  Score 30;  DB 1;  Length 404;
Best Local Similarity  100.0%;  Pred. No. 4.8e-25;
Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
           ||||||||||||||||||||||||||||||
Db      23 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 52
```

```
RESULT   2
RAGE_MOUSE
ID   RAGE_MOUSE     STANDARD;      PRT;   403 AA.
AC   Q62151;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Advanced glycosylation end product-specific receptor precursor
DE   (Receptor for advanced glycosylation end products).
GN   AGER OR RAGE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Lung;
RX   MEDLINE=97368045; PubMed=9224812;
RA   Renard C., Chappey O., Wautier M.P., Nagashima M., Lundh E.,
RA   Morser J., Zhao L., Schmidt A.M., Scherrmann J.M., Wautier J.L.;
RT   "Recombinant advanced glycation end product receptor pharmacokinetics
RT   in normal and diabetic rats.";
RL   Mol. Pharmacol. 52:54-62(1997).
CC   -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
CC       PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
CC       WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
CC       RATE IN DIABETES.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L33412; AAA40040.1; -.
DR   MGD; MGI:893592; Ager.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT   SIGNAL       1    22       POTENTIAL.
FT   CHAIN       23   403       ADVANCED GLYCOSYLATION END PRODUCT-
FT                              SPECIFIC RECEPTOR.
FT   DOMAIN      23   341       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   342   362       POTENTIAL.
FT   DOMAIN     363   403       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      31   105       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN     136   213       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN     251   307       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID    38    98       POTENTIAL.
FT   DISULFID   143   207       POTENTIAL.
FT   DISULFID   258   300       POTENTIAL.
FT   CARBOHYD    25    25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    80    80       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   403 AA;  42668 MW;  1279796FD1579357 CRC64;
```

```
  Query Match          43.3%;  Score 13;  DB 1;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 9e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 QNITARIGEPLVL 14
          |||||||||||||
Db     24 QNITARIGEPLVL 36


RESULT   3
RAGE_RAT
ID   RAGE_RAT       STANDARD;      PRT;   402 AA.
AC   Q63495;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Advanced glycosylation end product-specific receptor precursor
DE   (Receptor for advanced glycosylation end products).
GN   AGER OR RAGE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;
RX   MEDLINE=97368045; PubMed=9224812;
RA   Renard C., Chappey O., Wautier M.P., Nagashima M., Lundh E.,
RA   Morser J., Zhao L., Schmidt A.M., Scherrmann J.M., Wautier J.L.;
RT   "Recombinant advanced glycation end product receptor pharmacokinetics
RT   in normal and diabetic rats.";
RL   Mol. Pharmacol. 52:54-62(1997).
CC   -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
CC       PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
CC       WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
CC       RATE IN DIABETES.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L33413; AAA42027.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    402       ADVANCED GLYCOSYLATION END PRODUCT-
FT                                SPECIFIC RECEPTOR.
FT   DOMAIN       23    341       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    342    362       POTENTIAL.
FT   DOMAIN      363    402       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       31    105       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      136    212       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      250    306       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID     38     98       POTENTIAL.
FT   DISULFID    143    206       POTENTIAL.
FT   DISULFID    257    299       POTENTIAL.
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     80     80       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   402 AA;  42663 MW;  594481BC3A51E94E CRC64;


  Query Match          36.7%;  Score 11;  DB 1;  Length 402;
  Best Local Similarity  100.0%;  Pred. No. 0.00013;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 QNITARIGEPL 12
          ||||||||||
Db     24 QNITARIGEPL 34


RESULT   4
RAGE_BOVIN
ID   RAGE_BOVIN     STANDARD;      PRT;   416 AA.
AC   Q28173;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Advanced glycosylation end product-specific receptor precursor
DE   (Receptor for advanced glycosylation end products).
GN   AGER OR RAGE.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lung;
RX   MEDLINE=92340547; PubMed=1378843;
RA   Neeper M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,
RA   Elliston K., Stern D., Shaw A.;
RT   "Cloning and expression of a cell surface receptor for advanced
RT   glycosylation end products of proteins.";
RL   J. Biol. Chem. 267:14998-15004(1992).
CC   -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END
CC       PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS
CC       WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED
CC       RATE IN DIABETES.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M91212; AAA03575.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    416       ADVANCED GLYCOSYLATION END PRODUCT-
FT                                SPECIFIC RECEPTOR.
FT   DOMAIN       23    352       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    353    373       POTENTIAL.
FT   DOMAIN      374    416       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       31    105       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      136    213       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      262    318       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID     38     98       POTENTIAL.
```

```
FT   DISULFID    143    207       POTENTIAL.
FT   DISULFID    269    311       POTENTIAL.
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     80     80       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DOMAIN      391    396       POLY-GLU.
SQ   SEQUENCE   416 AA;  44182 MW;  B703815573E767AE CRC64;


  Query Match             33.3%;  Score 10;  DB 1;  Length 416;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     16 CKGAPKKPPQ 25
          ||||||||||
Db     38 CKGAPKKPPQ 47


RESULT   5
MUTH_HAEIN
ID   MUTH_HAEIN     STANDARD;      PRT;   223 AA.
AC   P44688;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA mismatch repair protein mutH.
GN   MUTH OR HI0403.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: SEQUENCE-SPECIFIC ENDONUCLEASE THAT CLEAVES UNMETHYLATED
CC       GATC SEQUENCES. IT IS INVOLVED IN DNA MISMATCH REPAIR (BY
CC       SIMILARITY).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32723; AAC22062.1; -.
DR   HSSP; P06722; 2AZO.
DR   TIGR; HI0403; -.
DR   InterPro; IPR004230; MutH.
DR   Pfam; PF02976; MutH; 1.
KW   DNA repair; Hydrolase; Endonuclease; Complete proteome.
SQ   SEQUENCE   223 AA;  24906 MW;  339A4EF9DA0E622A CRC64;


  Query Match             23.3%;  Score 7;  DB 1;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 ITARIGE 10
          |||||||
Db    165 ITARIGE 171


RESULT   6
VGLF_PVM
ID   VGLF_PVM       STANDARD;      PRT;   537 AA.
AC   P35949;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Pneumonia virus of mice (PVM).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX   NCBI_TaxID=11263;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92333256; PubMed=1629698;
RA   Chambers P., Pringle C.R., Easton A.J.;
RT   "Sequence analysis of the gene encoding the fusion glycoprotein of
RT   pneumonia virus of mice suggests possible conserved secondary
RT   structure elements in paramyxovirus fusion glycoproteins.";
RL   J. Gen. Virol. 73:1717-1724(1992).
CC   -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC       MEMBRANES, RESULTING IN VIRAL PENETRATION, & CAN DIRECT FUSION OF
CC       INFECTED CELLS WITH ADJOINING CELLS, RESULTING IN THE FORMATION OF
CC       SYNCYTIA.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D11128; BAA01902.1; -.
DR   PIR; JQ1619; JQ1619.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    537       FUSION GLYCOPROTEIN F0.
FT   CHAIN        23    101       F2 PROTEIN.
FT   CHAIN       102    537       F1 PROTEIN.
FT   TRANSMEM    491    514       POTENTIAL.
FT   CARBOHYD    463    463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    488    488       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   537 AA;  59366 MW;  BA6116EE2FABE702 CRC64;


  Query Match             23.3%;  Score 7;  DB 1;  Length 537;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 GEPLVLK 15
          |||||||
Db    434 GEPLVLK 440


RESULT   7
YG4L_YEAST
ID   YG4L_YEAST     STANDARD;      PRT;   110 AA.
AC   P53305;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
```

```
DE   Hypothetical 12.4 kDa protein in NAB1A-GPI1 intergenic region.
GN   YGR215W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=97435481; PubMed=9290212;
RA   Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT   "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT   chromosome VII.";
RL   Yeast 13:1077-1090(1997).
CC   -!- SIMILARITY: TO S.POMBE SPBC30D10.12C.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z73000; CAA97242.1; -.
DR   SGD; S0003447; YGR215W.
KW   Hypothetical protein.
SQ   SEQUENCE   110 AA;  12393 MW;  05008CA4F5D09004 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity   100.0%;  Pred. No. 9.6;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKK 22
           ||||||
Db      88 KGAPKK 93


RESULT   8
RS9_RICCN
ID   RS9_RICCN      STANDARD;      PRT;   159 AA.
AC   Q92IV4;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S9.
GN   RPSI OR RC0316.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
CC   -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008597; AAL02854.1; -.
DR   PROSITE; PS00360; RIBOSOMAL_S9; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   159 AA;  17967 MW;  03FE54B5529DC376 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 159;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 EPLVLK 15
           ||||||
Db      18 EPLVLK 23


RESULT   9
YCF4_GUITH
ID   YCF4_GUITH     STANDARD;      PRT;   181 AA.
AC   O78467;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Photosystem I assembly protein ycf4.
GN   YCF4.
OS   Guillardia theta (Cryptomonas phi).
OG   Chloroplast.
OC   Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX   NCBI_TaxID=55529;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99128221; PubMed=9929392;
RA   Douglas S.E., Penny S.L.;
RT   "The plastid genome of the cryptophyte alga, Guillardia theta:
RT   complete sequence and conserved synteny groups confirm its common
RT   ancestry with red algae.";
RL   J. Mol. Evol. 48:236-244(1999).
CC   -!- FUNCTION: Seems to be required for the assembly of the photosystem
CC       I complex (By similarity).
CC   -!- SUBCELLULAR LOCATION: Thylakoid membrane-associated (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE YCF4 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041468; AAC35658.1; -.
DR   InterPro; IPR003359; Ycf4.
DR   Pfam; PF02392; Ycf4; 1.
KW   Photosynthesis; Thylakoid; Transmembrane; Chloroplast.
FT   TRANSMEM     19     41       POTENTIAL.
FT   TRANSMEM     61     83       POTENTIAL.
SQ   SEQUENCE   181 AA;  20921 MW;  9B74EC86AFEC6F48 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 181;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 RIGEPL 12
            ||||||
Db     153 RIGEPL 158


RESULT  10
YA79_HAEIN
ID   YA79_HAEIN     STANDARD;      PRT;   210 AA.
AC   P45023;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
```

```
DE   Hypothetical amino-acid ABC transporter permease protein HI0179.
GN   HI1079.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RD / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus
RT   influenzae Rd.";
RL   Science 269:496-512(1995).
RN   [2]
RP   REVISIONS.
RA   White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
CC       SYSTEM FOR AN AMINO-ACID. PROBABLY RESPONSIBLE FOR THE
CC       TRANSLOCATION OF THE SUBSTRATE ACROSS THE MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
CC   -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC       PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE HISMQ
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32788; AAC22735.1; -.
DR   TIGR; HI1079; -.
DR   InterPro; IPR000515; BPD_transp.
DR   Pfam; PF00528; BPD_transp; 1.
DR   PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; 1.
KW   Hypothetical protein; Transport; Amino-acid transport; Transmembrane;
KW   Inner membrane; Complete proteome.
FT   TRANSMEM     10     30       POTENTIAL.
FT   TRANSMEM     57     77       POTENTIAL.
FT   TRANSMEM     79     99       POTENTIAL.
FT   TRANSMEM    177    197       POTENTIAL.
SQ   SEQUENCE   210 AA;  23393 MW;  662C7C590DC4FDCD CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITA 6
          ||||||
Db    167 AQNITA 172


RESULT  11
VGP8_EBV
ID   VGP8_EBV       STANDARD;      PRT;   234 AA.
AC   P03224;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   Probable membrane antigen GP85.
GN   BDLF3.
OS   Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84270667; PubMed=6087149;
RA   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA   Tuffnell P.S., Barrell B.G.;
RT   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL   Nature 310:207-211(1984).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01555; CAA24835.1; -.
DR   PIR; A03787; QQBE43.
DR   PIR; S33040; S33040.
KW   Membrane; Glycoprotein; Late protein.
SQ   SEQUENCE   234 AA;  23791 MW;  48D6CED09331119E CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITA 6
          ||||||
Db    108 AQNITA 113


RESULT  12
E434_ADECT
ID   E434_ADECT     STANDARD;      PRT;   259 AA.
AC   P87568;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Early E4 30 kDa protein.
OS   Canine adenovirus type 2 (strain Toronto A 26-61).
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=69152;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Campbell J.B., Zhao Y.;
RT   "Complete DNA sequence and genomic organization of canine
RT   adenovirus type 2.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE ADENOVIRUS E4 30 TO 34 kDa PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U77082; AAB38735.1; -.
KW   Early protein.
SQ   SEQUENCE   259 AA;  30014 MW;  9C966CA011C2A745 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 259;
```

```
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LVLKCK 17
           ||||||
Db     178 LVLKCK 183


RESULT  13
E434_ADECC
ID   E434_ADECC     STANDARD;      PRT;   265 AA.
AC   Q65962;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Early E4 31 kDa protein.
OS   Canine adenovirus type 1 (strain CLL).
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=69150;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Campbell J.B., Zhao Y.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 248-265 FROM N.A.
RX   MEDLINE=91272490; PubMed=1828920;
RA   Dragulev B.P., Sira S., Abouhaidar M.G., Campbell J.B.;
RT   "Sequence analysis of putative E3 and fiber genomic regions of two
RT   strains of canine adenovirus type 1.";
RL   Virology 183:298-305(1991).
CC   -!- SIMILARITY: BELONGS TO THE ADENOVIRUS E4 30 TO 34 kDa PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U55001; AAB05452.1; -.
KW   Early protein.
SQ   SEQUENCE   265 AA;  30822 MW;  92606E5E30134103 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 265;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LVLKCK 17
           ||||||
Db     184 LVLKCK 189


RESULT  14
E434_ADECR
ID   E434_ADECR     STANDARD;      PRT;   265 AA.
AC   Q96690;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Early E4 31 kDa protein.
OS   Canine adenovirus type 1 (strain RI261).
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=69151;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97275900; PubMed=9129661;
RA   Morrison M.D., Onions D.E., Nicolson L.;
RT   "Complete DNA sequence of canine adenovirus type 1.";
RL   J. Gen. Virol. 78:873-878(1997).
CC   -!- SIMILARITY: BELONGS TO THE ADENOVIRUS E4 31 TO 34 kDa PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y07760; CAA69046.1; -.
KW   Early protein.
SQ   SEQUENCE   265 AA;  30737 MW;  14B3DC9DCB1D2E15 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 265;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LVLKCK 17
           ||||||
Db     184 LVLKCK 189


RESULT  15
MIND_GUITH
ID   MIND_GUITH     STANDARD;      PRT;   269 AA.
AC   O78436;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative septum site-determining protein minD.
GN   MIND.
OS   Guillardia theta (Cryptomonas phi).
OG   Chloroplast.
OC   Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX   NCBI_TaxID=55529;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99128221; PubMed=9929392;
RA   Douglas S.E., Penny S.L.;
RT   "The plastid genome of the cryptophyte alga, Guillardia theta:
RT   complete sequence and conserved synteny groups confirm its common
RT   ancestry with red algae.";
RL   J. Mol. Evol. 48:236-244(1999).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041468; AAC35621.1; -.
DR   InterPro; IPR000707; ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Chloroplast.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   269 AA;  29455 MW;  BC363E954E689EA2 CRC64;


  Query Match             20.0%;  Score 6;  DB 1;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 GEPLVL 14
           ||||||
Db     215 GEPLVL 220
```

Search completed: July 31, 2002, 15:08:15
Job time: 220 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:04:10 ; Search time 25.08 Seconds
(without alignments)
206.932 Million cell updates/sec

Title: US-08-948-131-1
Perfect score: 30
Sequence: 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13	43.3	402	11	O35444	O35444 mus musculu
2	9	30.0	32	6	Q9TRQ1	Q9trq1 bos taurus
3	7	23.3	237	5	Q9V5N7	Q9v5n7 drosophila
4	7	23.3	330	16	Q9K1F6	Q9k1f6 neisseria m
5	7	23.3	330	16	Q9JX44	Q9jx44 neisseria m
6	7	23.3	595	10	Q9LQP6	Q9lqp6 arabidopsis
7	7	23.3	742	5	Q9W322	Q9w322 drosophila
8	6	20.0	61	5	Q9NFJ5	Q9nfj5 trypanosoma
9	6	20.0	71	5	Q9N6K0	Q9n6k0 trypanosoma
10	6	20.0	76	5	Q9N6I0	Q9n6i0 trypanosoma
11	6	20.0	96	10	O80389	O80389 cucumis sat
12	6	20.0	99	10	Q9FR57	Q9fr57 lycopersico
13	6	20.0	101	12	Q9JH47	Q9jh47 human papil
14	6	20.0	127	15	Q9EAJ5	Q9eaj5 human immun
15	6	20.0	155	10	Q9SCQ3	Q9scq3 arabidopsis
16	6	20.0	159	16	Q92IV4	Q92iv4 rickettsia
17	6	20.0	204	16	Q92TQ0	Q92tq0 rhizobium m
18	6	20.0	209	16	O06630	O06630 mycobacteri
19	6	20.0	224	16	Q92JS5	Q92js5 rhizobium m
20	6	20.0	230	16	Q9RRE2	Q9rre2 deinococcus
21	6	20.0	231	16	Q92KL9	Q92kl9 rhizobium m
22	6	20.0	232	14	Q91UW4	Q91uw4 plasmid pip
23	6	20.0	257	16	O06157	O06157 mycobacteri
24	6	20.0	259	16	Q9CCU3	Q9ccu3 mycobacteri
25	6	20.0	261	2	O07711	O07711 mycobacteri
26	6	20.0	271	16	O67348	O67348 aquifex aeo
27	6	20.0	298	10	Q9ZS04	Q9zs04 arabidopsis
28	6	20.0	298	16	Q9A7U8	Q9a7u8 caulobacter
29	6	20.0	318	16	Q987Q7	Q987q7 rhizobium l
30	6	20.0	326	16	Q9HXX1	Q9hxx1 pseudomonas
31	6	20.0	334	5	Q9V6W0	Q9v6w0 drosophila
32	6	20.0	334	16	Q9HXF7	Q9hxf7 pseudomonas
33	6	20.0	339	11	Q9QYV4	Q9qyv4 rattus norv
34	6	20.0	339	11	Q62708	Q62708 rattus norv
35	6	20.0	359	2	O88036	O88036 streptomyce
36	6	20.0	359	8	Q9TJR6	Q9tjr6 prototheca
37	6	20.0	377	16	Q97IK0	Q97ik0 clostridium
38	6	20.0	387	10	Q9XIS0	Q9xis0 arabidopsis
39	6	20.0	389	9	Q38644	Q38644 bacteriopha
40	6	20.0	392	2	Q9FCL0	Q9fcl0 streptomyce
41	6	20.0	411	10	Q9M3H9	Q9m3h9 cicer ariet
42	6	20.0	426	5	Q9I7W7	Q9i7w7 drosophila
43	6	20.0	433	2	Q52912	Q52912 rhizobium m
44	6	20.0	433	16	Q92TA1	Q92ta1 rhizobium m
45	6	20.0	452	10	Q9LRM5	Q9lrm5 arabidopsis

ALIGNMENTS

RESULT 1
O35444
ID O35444 PRELIMINARY; PRT; 402 AA.
AC O35444;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RAGE.
GN RAGE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,
RA Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.;
RT "Sequence of the mouse major histocompatibility locus class III
RT region.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF030001; AAB82007.1; -.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Immunoglobulin domain.
SQ SEQUENCE 402 AA; 42653 MW; DBFDC50A6C8CB902 CRC64;

Query Match 43.3%; Score 13; DB 11; Length 402;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QNITARIGEPLVL 14
||||||||||||
Db 24 QNITARIGEPLVL 36

```
RESULT   2
Q9TRQ1
ID   Q9TRQ1      PRELIMINARY;      PRT;    32 AA.
AC   Q9TRQ1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   35 KDA ADVANCED GLYCOSYLATION END PRODUCT BINDING PROTEIN (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92340546; PubMed=1321822;
RA   Schmidt A.M., Vianna M., Gerlach M., Brett J., Ryan J., Kao J.,
RA   Esposito C., Hegarty H., Hurley W., Clauss M.;
RT   "Isolation and characterization of two binding proteins for advanced
RT   glycosylation end products from bovine lung which are present on the
RT   endothelial cell surface.";
RL   J. Biol. Chem. 267:14987-14997(1992).
SQ   SEQUENCE   32 AA;  3507 MW;  AE4C3147CE5B3D91 CRC64;


  Query Match             30.0%;  Score 9;  DB 6;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 0.0074;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKKPPQ 25
           |||||||||
Db      17 KGAPKKPPQ 25


RESULT   3
Q9V5N7
ID   Q9V5N7      PRELIMINARY;      PRT;   237 AA.
AC   Q9V5N7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG12934 PROTEIN.
GN   CG12934.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003828; AAF58764.1; -.
DR   FlyBase; FBgn0033541; CG12934.
SQ   SEQUENCE   237 AA;  27208 MW;  C050101AD183FAF7 CRC64;


  Query Match             23.3%;  Score 7;  DB 5;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 6.5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      22 KPPQRLE 28
           |||||||
Db     106 KPPQRLE 112


RESULT   4
Q9K1F6
ID   Q9K1F6      PRELIMINARY;      PRT;   330 AA.
AC   Q9K1F6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE C.
GN   NMB0198.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / SEROGROUP B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002377; AAF40655.1; -.
DR   TIGR; NMB0198; -.
DR   InterPro; IPR000613; PseudoU_synth.
DR   InterPro; IPR002990; PSI_RLU.
DR   InterPro; IPR002942; S4.
DR   Pfam; PF00849; PseudoU_synth_2; 1.
DR   ProDom; PD001819; PseudoU_synth; 1.
DR   SMART; SM00363; S4; 1.
DR   PROSITE; PS01129; PSI_RLU; 1.
KW   Complete proteome.
SQ   SEQUENCE   330 AA;  36682 MW;  F2058C52ACE443EC CRC64;
```

```
  Query Match           23.3%;  Score 7;  DB 16;  Length 330;
  Best Local Similarity  100.0%;  Pred. No. 8.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 GEPLVLK 15
          |||||||
Db    304 GEPLVLK 310


RESULT   5
Q9JX44
ID   Q9JX44      PRELIMINARY;      PRT;   330 AA.
AC   Q9JX44;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update).
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE C (EC 4.2.1.70).
GN   RLUC OR NMA0070.
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162752; CAB83386.1; -.
DR   InterPro; IPR000613; PseudoU_synth.
DR   InterPro; IPR002990; PSI_RLU.
DR   InterPro; IPR002942; S4.
DR   Pfam; PF00849; PseudoU_synth_2; 1.
DR   ProDom; PD001819; PseudoU_synth; 1.
DR   SMART; SM00363; S4; 1.
DR   PROSITE; PS01129; PSI_RLU; 1.
KW   Lyase; Complete proteome.
SQ   SEQUENCE   330 AA;  36768 MW;  9B1AB94890F675EA CRC64;


  Query Match           23.3%;  Score 7;  DB 16;  Length 330;
  Best Local Similarity  100.0%;  Pred. No. 8.6;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 GEPLVLK 15
          |||||||
Db    304 GEPLVLK 310


RESULT   6
Q9LQP6
ID   Q9LQP6      PRELIMINARY;      PRT;   595 AA.
AC   Q9LQP6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   F24B9.20.
GN   F24B9.20.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Liu S., Yu G., Sakano H., Jhaveri A., Lee J., Lenz C., Pham P.,
RA   Toriumi M., Chin C., Choi E., Chiou J., Gonzalez A., Chung M.,
RA   Howng B., Koo T., Li J., Liu A., Vaysberg M., Altafi H., Brooks S.,
RA   Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.,
RA   Johnson-Hopson C., Khan S., Kim C., Lam B., Nguyen M., Palm C.,
RA   Shinn P., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT   "The sequence of BAC F24B9 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Theologis A.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007583; AAF75084.1; -.
SQ   SEQUENCE   595 AA;  63982 MW;  FF5B15F9BB28B87E CRC64;


  Query Match           23.3%;  Score 7;  DB 10;  Length 595;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 RIGEPLV 13
          |||||||
Db    550 RIGEPLV 556


RESULT   7
Q9W322
ID   Q9W322      PRELIMINARY;      PRT;   742 AA.
AC   Q9W322;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CG3106 PROTEIN (GH10201P).
GN   CG3106.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
```

```
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Y, CN BW SP;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003448; AAF46515.1; -.
DR   EMBL; AY058329; AAL13558.1; -.
DR   FlyBase; FBgn0030148; CG3106.
DR   InterPro; IPR003862; UPF0063.
DR   Pfam; PF02712; UPF0063; 1.
SQ   SEQUENCE   742 AA;  84476 MW;  EFF31DF7DC6DE02D CRC64;


  Query Match           23.3%;  Score 7;  DB 5;  Length 742;
  Best Local Similarity  100.0%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKKP 23
           |||||||
Db     254 KGAPKKP 260


RESULT   8
Q9NFJ5
ID   Q9NFJ5      PRELIMINARY;      PRT;    61 AA.
AC   Q9NFJ5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HISTONE H1.
GN   H1A61.
OS   Trypanosoma brucei brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TREU 927/4;
RA   Grueter E.;
RL   Thesis (2000), Department of Parasitology, Institute of Zoology,
RL   SWITZERLAND.
DR   EMBL; AJ287603; CAB76188.1; -.
SQ   SEQUENCE   61 AA;  6012 MW;  7B66DA10A7D33B93 CRC64;


  Query Match           20.0%;  Score 6;  DB 5;  Length 61;
  Best Local Similarity  100.0%;  Pred. No. 26;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKK 22
           ||||||
Db      30 KGAPKK 35


RESULT   9
Q9N6K0
ID   Q9N6K0      PRELIMINARY;      PRT;    71 AA.
AC   Q9N6K0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HISTONE H1.
GN   H1E71 OR H1A71.
OS   Trypanosoma brucei brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TREU 927/4;
RA   Grueter E.;
RL   Thesis (2000), Department of Parasitology, Institute of Zoology,.
DR   EMBL; AJ400881; CAB90839.1; -.
DR   EMBL; AJ287600; CAB76185.1; -.
SQ   SEQUENCE   71 AA;  7005 MW;  478C7B6994847C80 CRC64;


  Query Match           20.0%;  Score 6;  DB 5;  Length 71;
  Best Local Similarity  100.0%;  Pred. No. 29;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKK 22
           ||||||
Db      30 KGAPKK 35


RESULT  10
Q9N6I0
ID   Q9N6I0      PRELIMINARY;      PRT;    76 AA.
AC   Q9N6I0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE H1.
GN   H1E76 OR H1F76.
OS   Trypanosoma brucei brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TREU 927/4;
RA   Grueter E.;
RL   Thesis (2000), Department of Parasitology, Institute of Zoology,
RL   SWITZERLAND.
DR   EMBL; AJ287605; CAB76190.1; -.
DR   EMBL; AJ287601; CAB76186.1; -.
DR   InterPro; IPR001386; Linker_histone.
DR   PRINTS; PR00624; HISTONEH5.
SQ   SEQUENCE   76 AA;  7573 MW;  6079EF4982B50767 CRC64;


  Query Match           20.0%;  Score 6;  DB 5;  Length 76;
  Best Local Similarity  100.0%;  Pred. No. 31;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 KGAPKK 22
           ||||||
Db      30 KGAPKK 35


RESULT  11
O80389
ID   O80389      PRELIMINARY;      PRT;    96 AA.
AC   O80389;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CYSTEIN PROTEINASE INHIBITOR.
OS   Cucumis sativus (Cucumber).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX   NCBI_TaxID=3659;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamakawa S.;
```

```
RT   "Cystein proteinase inhibitor.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB014760; BAA28867.1; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C_M.
DR   Pfam; PF00031; cystatin; 1.
DR   ProDom; PD001231; Cystatin_C_M; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; UNKNOWN_1.
SQ   SEQUENCE   96 AA;  10973 MW;  027252E14BB6C4F4 CRC64;


  Query Match           20.0%;  Score 6;  DB 10;  Length 96;
  Best Local Similarity  100.0%;  Pred. No. 38;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     12 LVLKCK 17
          ||||||
Db     59 LVLKCK 64


RESULT  12
Q9FR57
ID   Q9FR57      PRELIMINARY;      PRT;    99 AA.
AC   Q9FR57;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SELF-PRUNING INTERACTING PROTEIN 1.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. VFNT;
RA   Pnueli L., Gutfinger T., Hareven D., Ben-Naim O., Ron N., Adir N.,
RA   Lifschitz E.;
RT   "Tomato SP-interacting proteins define a conserved signaling system
RT   that regulates shoot architecture and flowering.";
RL   Plant Cell 0:0-0(2001).
DR   EMBL; AF175963; AAG43410.1; -.
SQ   SEQUENCE   99 AA;  11408 MW;  E55F6975B7BA97A3 CRC64;


  Query Match           20.0%;  Score 6;  DB 10;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 39;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     18 GAPKKP 23
          ||||||
Db     45 GAPKKP 50


RESULT  13
Q9JH47
ID   Q9JH47      PRELIMINARY;      PRT;   101 AA.
AC   Q9JH47;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   START CODON IS NOT IDENTIFIED (FRAGMENT).
GN   E4.
OS   Human papillomavirus type 69.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=37121;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Matsukura T., Sata T.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20087389; PubMed=10618284;
RA   Kino N., Sata T., Sato Y., Sugase M., Matsukura T.;
RT   "Molecular cloning and nucleotide sequence analysis of a novel human
RT   papillomavirus (type 82) associated with vaginal intraepithelial
RT   neoplasia.";
RL   Clin. Diagn. Lab. Immunol. 7:91-95(2000).
DR   EMBL; AB027020; BAA90731.1; -.
DR   InterPro; IPR003861; Papilloma_E4.
DR   Pfam; PF02711; Pap_E4; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   101 AA;  11267 MW;  32899BE169CAC792 CRC64;


  Query Match           20.0%;  Score 6;  DB 12;  Length 101;
  Best Local Similarity  100.0%;  Pred. No. 40;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     18 GAPKKP 23
          ||||||
Db     39 GAPKKP 44


RESULT  14
Q9EAJ5
ID   Q9EAJ5      PRELIMINARY;      PRT;   127 AA.
AC   Q9EAJ5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROTEASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=99-117635;
RX   MEDLINE=21178648; PubMed=11282016;
RA   Barlow K.L., Tatt I.D., Cane P.A., Pillay D., Clewley J.P.;
RT   " Recombinant strains of HIV Type 1 in the United Kingdom.";
RL   AIDS Res. Hum. Retroviruses 17:467-474(2001).
CC   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC       KNOWN AS THE RETROPEPSIN FAMILY.
DR   EMBL; AJ296662; CAC03693.1; -.
DR   HSSP; P04584; 1PHV.
DR   InterPro; IPR001969; Asp_protease.
DR   InterPro; IPR001995; Asp_prot_retrov.
DR   Pfam; PF00077; rvp; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
DR   PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW   Aspartyl protease; Hydrolase.
FT   NON_TER       1       1
FT   NON_TER     127     127
SQ   SEQUENCE   127 AA;  13648 MW;  CEAFF2745863D5F4 CRC64;


  Query Match           20.0%;  Score 6;  DB 15;  Length 127;
  Best Local Similarity  100.0%;  Pred. No. 48;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 ITARIG 9
          ||||||
Db     39 ITARIG 44


RESULT  15
Q9SCQ3
ID   Q9SCQ3      PRELIMINARY;      PRT;   155 AA.
AC   Q9SCQ3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   HYPOTHETICAL 17.0 KDA PROTEIN.
GN   T4D2.10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; 'core 'eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA   Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL132958; CAB64211.1; -.
DR   InterPro; IPR000922; Gal_lectin.
DR   Pfam; PF02140; Gal_Lectin; 1.
DR   ProDom; PD005612; Gal_lectin; 1.
DR   PROSITE; PS50228; SUEL_LECTIN; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   155 AA;  16990 MW;  815302F74AE2EFCB CRC64;


  Query Match             20.0%;  Score 6;  DB 10;  Length 155;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 KCKGAP 20
           ||||||
Db     139 KCKGAP 144


Search completed: July 31, 2002, 15:08:00
Job time: 230 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 31, 2002, 15:03:45 ; Search time 12.99 Seconds
(without alignments)
56.410 Million cell updates/sec

Title: US-08-948-131-1
Perfect score: 30
Sequence: 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	318	2	US-08-633-148-4	Sequence 4, Appli
2	30	100.0	340	2	US-08-633-148-2	Sequence 2, Appli
3	23	76.7	278	2	US-08-432-016-5	Sequence 5, Appli
4	23	76.7	278	2	US-08-684-594-5	Sequence 5, Appli
5	16	53.3	16	2	US-08-633-148-18	Sequence 18, Appl
6	15	50.0	15	2	US-08-633-148-12	Sequence 12, Appl
7	10	33.3	10	2	US-08-633-148-5	Sequence 5, Appli
8	10	33.3	10	2	US-08-633-148-13	Sequence 13, Appl
9	6	20.0	10	4	US-08-685-558A-2	Sequence 2, Appli
10	6	20.0	479	1	US-08-416-788-2	Sequence 2, Appli
11	6	20.0	481	1	US-08-173-436A-2	Sequence 2, Appli
12	6	20.0	481	2	US-08-748-485-8	Sequence 8, Appli
13	6	20.0	513	4	US-08-685-558A-18	Sequence 18, Appl
14	6	20.0	741	1	US-08-277-231A-4	Sequence 4, Appli
15	6	20.0	741	2	US-08-473-750-7	Sequence 7, Appli
16	6	20.0	741	2	US-08-477-326-7	Sequence 7, Appli
17	6	20.0	1103	3	US-09-162-373-1	Sequence 1, Appli
18	6	20.0	1103	4	US-09-467-946-1	Sequence 1, Appli
19	6	20.0	1958	1	US-07-945-283-2	Sequence 2, Appli
20	5	16.7	11	1	US-08-336-343A-23	Sequence 23, Appl
21	5	16.7	14	1	US-08-137-117D-153	Sequence 153, App
22	5	16.7	14	1	US-08-137-117D-155	Sequence 155, App
23	5	16.7	14	1	US-08-137-117D-158	Sequence 158, App
24	5	16.7	14	2	US-08-436-717-153	Sequence 153, App
25	5	16.7	14	2	US-08-436-717-155	Sequence 155, App
26	5	16.7	14	2	US-08-436-717-158	Sequence 158, App
27	5	16.7	20	4	US-08-602-999A-134	Sequence 134, App
28	5	16.7	20	4	US-09-242-131A-1	Sequence 1, Appli
29	5	16.7	20	4	US-09-615-283-1	Sequence 1, Appli
30	5	16.7	22	4	US-09-242-131A-7	Sequence 7, Appli
31	5	16.7	22	4	US-09-615-283-7	Sequence 7, Appli
32	5	16.7	24	4	US-08-525-539A-22	Sequence 22, Appl
33	5	16.7	26	2	US-08-637-759B-326	Sequence 326, App
34	5	16.7	26	3	US-08-871-355A-326	Sequence 326, App
35	5	16.7	26	4	US-08-525-539A-20	Sequence 20, Appl
36	5	16.7	26	4	US-09-201-945-326	Sequence 326, App
37	5	16.7	49	4	US-09-085-305-22	Sequence 22, Appl
38	5	16.7	59	4	US-09-085-305-21	Sequence 21, Appl
39	5	16.7	63	2	US-08-637-759B-191	Sequence 191, App
40	5	16.7	63	3	US-08-871-355A-191	Sequence 191, App
41	5	16.7	63	4	US-09-201-945-191	Sequence 191, App
42	5	16.7	72	4	US-09-188-930-153	Sequence 153, App
43	5	16.7	72	4	US-09-188-930-304	Sequence 304, App
44	5	16.7	75	4	US-09-450-072-70	Sequence 70, Appl
45	5	16.7	80	4	US-09-085-305-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT 1
US-08-633-148-4
; Sequence 4, Application US/08633148
; Patent No. 5864018
; GENERAL INFORMATION:
;   APPLICANT: MORSER, MICHAEL J.
;   APPLICANT: NAGASHIMA, MARIKO
;   APPLICANT: HOLLANDER, DORIS A.
;   TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
;   TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: TOWNSEND & TOWNSENT & CREW LLP
;     STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
;     CITY: SAN FRANCISCO
;     STATE: CALIFORNIA
;     COUNTRY: U.S.A.
;     ZIP: 94111
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/633,148
;     FILING DATE: 16-APR-1996
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: MURPHY ESQ., MATTHEW B.
;     REGISTRATION NUMBER: 39,787
;     REFERENCE/DOCKET NUMBER: 014618-005600US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 326-2400
;     TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 318 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-633-148-4

  Query Match 100.0%; Score 30; DB 2; Length 318;
  Best Local Similarity 100.0%; Pred. No. 1.3e-24;
  Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
     ||||||||||||||||||||||||||||||
```

```
Db      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30


RESULT   2
US-08-633-148-2
; Sequence 2, Application US/08633148
; Patent No. 5864018
;  GENERAL INFORMATION:
;    APPLICANT:  MORSER, MICHAEL J.
;    APPLICANT:  NAGASHIMA, MARIKO
;    APPLICANT:  HOLLANDER, DORIS A.
;    TITLE OF INVENTION:  ANTIBODIES TO ADVANCED GLYCOSYLATION
;    TITLE OF INVENTION:  END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND & TOWNSENT & CREW LLP
;      STREET:  TWO EMBARCADERO CENTER, 8TH FLOOR
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/633,148
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY ESQ., MATTHEW B.
;      REGISTRATION NUMBER:  39,787
;      REFERENCE/DOCKET NUMBER:  014618-005600US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  340 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-633-148-2


  Query Match           100.0%;  Score 30;  DB 2;  Length 340;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-24;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 30
          ||||||||||||||||||||||||||||||
Db     23 AQNITARIGEPLVLKCKGAPKKPPQRLEWK 52


RESULT   3
US-08-432-016-5
; Sequence 5, Application US/08432016
; Patent No. 5968768
;  GENERAL INFORMATION:
;    APPLICANT:  HAYNES, BARTON F.
;    APPLICANT:  ARUFFO, ALEJANDRO
;    APPLICANT:  PATEL, DHAVALKUMAR
;    APPLICANT:  BOWEN, MICHAEL A.
;    APPLICANT:  MARQUARDT, HANS
;    TITLE OF INVENTION:  CD6 LIGAND
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/432,016
;      FILING DATE:  01-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,350
;      FILING DATE:  02-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/143,903
;      FILING DATE:  02-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WILSON, MARY J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  1579-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  278 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-432-016-5


  Query Match           76.7%;  Score 23;  DB 2;  Length 278;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-17;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 IGEPLVLKCKGAPKKPPQRLEWK 30
          |||||||||||||||||||||||
Db      1 IGEPLVLKCKGAPKKPPQRLEWK 23


RESULT   4
US-08-684-594-5
; Sequence 5, Application US/08684594
; Patent No. 5998172
;  GENERAL INFORMATION:
;    APPLICANT:  HAYNES, BARTON F.
;    APPLICANT:  ARUFFO, ALEJANDRO
;    APPLICANT:  PATEL, DHAVALKUMAR
;    APPLICANT:  BOWEN, MICHAEL A.
;    APPLICANT:  MARQUARDT, HANS
;    TITLE OF INVENTION:  CD6 LIGAND
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/684,594
;      FILING DATE:  18-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/432,016
```

; FILING DATE: 01-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/333,350
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/143,903
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-684-594-5

Query Match 76.7%; Score 23; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 3.1e-17;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 IGEPLVLKCKGAPKKPPQRLEWK 30
||||||||||||||||||||||||
Db 1 IGEPLVLKCKGAPKKPPQRLEWK 23

RESULT 5
US-08-633-148-18
; Sequence 18, Application US/08633148
; Patent No. 5864018
; GENERAL INFORMATION:
; APPLICANT: MORSER, MICHAEL J.
; APPLICANT: NAGASHIMA, MARIKO
; APPLICANT: HOLLANDER, DORIS A.
; TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
; TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND & TOWNSENT & CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,148
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY ESQ., MATTHEW B.
; REGISTRATION NUMBER: 39,787
; REFERENCE/DOCKET NUMBER: 014618-005600US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-633-148-18

Query Match 53.3%; Score 16; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.2e-11;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLKC 16
||||||||||||||||
Db 1 AQNITARIGEPLVLKC 16

RESULT 6
US-08-633-148-12
; Sequence 12, Application US/08633148
; Patent No. 5864018
; GENERAL INFORMATION:
; APPLICANT: MORSER, MICHAEL J.
; APPLICANT: NAGASHIMA, MARIKO
; APPLICANT: HOLLANDER, DORIS A.
; TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
; TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND & TOWNSENT & CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,148
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY ESQ., MATTHEW B.
; REGISTRATION NUMBER: 39,787
; REFERENCE/DOCKET NUMBER: 014618-005600US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-633-148-12

Query Match 50.0%; Score 15; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.9e-10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQNITARIGEPLVLK 15
|||||||||||||||
Db 1 AQNITARIGEPLVLK 15

RESULT 7
US-08-633-148-5
; Sequence 5, Application US/08633148
; Patent No. 5864018
; GENERAL INFORMATION:
; APPLICANT: MORSER, MICHAEL J.
; APPLICANT: NAGASHIMA, MARIKO

```
;    APPLICANT:  HOLLANDER, DORIS A.
;    TITLE OF INVENTION:  ANTIBODIES TO ADVANCED GLYCOSYLATION
;    TITLE OF INVENTION:  END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND & TOWNSENT & CREW LLP
;      STREET:  TWO EMBARCADERO CENTER, 8TH FLOOR
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/633,148
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY ESQ., MATTHEW B.
;      REGISTRATION NUMBER:  39,787
;      REFERENCE/DOCKET NUMBER:  014618-005600US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-633-148-5


  Query Match             33.3%;  Score 10;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 CKGAPKKPPQ 25
           ||||||||||
Db       1 CKGAPKKPPQ 10


RESULT   8
US-08-633-148-13
; Sequence 13, Application US/08633148
; Patent No. 5864018
;  GENERAL INFORMATION:
;    APPLICANT:  MORSER, MICHAEL J.
;    APPLICANT:  NAGASHIMA, MARIKO
;    APPLICANT:  HOLLANDER, DORIS A.
;    TITLE OF INVENTION:  ANTIBODIES TO ADVANCED GLYCOSYLATION
;    TITLE OF INVENTION:  END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND & TOWNSENT & CREW LLP
;      STREET:  TWO EMBARCADERO CENTER, 8TH FLOOR
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/633,148
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY ESQ., MATTHEW B.
;      REGISTRATION NUMBER:  39,787
;      REFERENCE/DOCKET NUMBER:  014618-005600US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-633-148-13


  Query Match             33.3%;  Score 10;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 CKGAPKKPPQ 25
           ||||||||||
Db       1 CKGAPKKPPQ 10


RESULT   9
US-08-685-558A-2
; Sequence 2, Application US/08685558A
; Patent No. 6225081
;  GENERAL INFORMATION:
;    APPLICANT:  SHIMOMURA, Takeshi
;    APPLICANT:  KAWAGUCHI, Toshiya
;    APPLICANT:  KITAMURA, Naomi
;    APPLICANT:  MIYAZAWA, Keiji
;    TITLE OF INVENTION:  NOVEL PROTEIN, DNA CODING FOR SAME
;    TITLE OF INVENTION:  AND METHOD OF PRODUCING THE PROTEIN
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,558A
;      FILING DATE:  24-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JPA Hei 7-187135
;      FILING DATE:  24-JUL-1995
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal fragment
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      STRAIN:  MKN45
US-08-685-558A-2


  Query Match             20.0%;  Score 6;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      10 EPLVLK 15
           ||||||
Db       5 EPLVLK 10


RESULT  10
US-08-416-788-2
; Sequence 2, Application US/08416788
; Patent No. 5780245
;  GENERAL INFORMATION:
;    APPLICANT:  Maroteaux, Luc
;    TITLE OF INVENTION:  No. 5780245el Polypeptides Having a Serotonin
;    TITLE OF INVENTION:  Receptor Activity, Nucleic Acids Coding for These
;    TITLE OF INVENTION:  Polypeptides and Uses
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;      STREET:  500 Arcola Road, 3C43
;      CITY:  Collegeville
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19426-0107
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,788
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR93/01012
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92-12280
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, Julie K..
;      REGISTRATION NUMBER:  38,619
;      REFERENCE/DOCKET NUMBER:  EX92008-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)454-3839
;      TELEFAX:  (610)454-3808
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 479 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-788-2


  Query Match             20.0%;  Score 6;  DB 1;  Length 479;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      22 KPPQRL 27
           ||||||
Db     253 KPPQRL 258


RESULT  11
US-08-173-436A-2
; Sequence 2, Application US/08173436A
; Patent No. 5698444
;  GENERAL INFORMATION:
;    APPLICANT:  Baez, Melvyn
;    APPLICANT:  Kursar, Jonathon D.
;    TITLE OF INVENTION:  SEROTONIN RECEPTOR PROTEIN AND RELATED
;    TITLE OF INVENTION:  NUCLEIC ACID COMPOUNDS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center/Patent Division
;      CITY:  Indianapolis
;      STATE:  IN
;      COUNTRY:  US
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,436A
;      FILING DATE:  23-DEC-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gaylo, Paul J.
;      REGISTRATION NUMBER:  36,808
;      REFERENCE/DOCKET NUMBER:  X-9367
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-0756
;      TELEFAX:  317-276-3861
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  481 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-173-436A-2


  Query Match             20.0%;  Score 6;  DB 1;  Length 481;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      22 KPPQRL 27
           ||||||
Db     254 KPPQRL 259


RESULT  12
US-08-748-485-8
; Sequence 8, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
```

```
;      REFERENCE/DOCKET NUMBER:  PF-0159 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  481 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  475198
US-08-748-485-8


  Query Match             20.0%;  Score 6;  DB 2;  Length 481;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      22 KPPQRL 27
           ||||||
Db     254 KPPQRL 259


RESULT  13
US-08-685-558A-18
; Sequence 18, Application US/08685558A
; Patent No. 6225081
;  GENERAL INFORMATION:
;    APPLICANT:  SHIMOMURA, Takeshi
;    APPLICANT:  KAWAGUCHI, Toshiya
;    APPLICANT:  KITAMURA, Naomi
;    APPLICANT:  MIYAZAWA, Keiji
;    TITLE OF INVENTION:  NOVEL PROTEIN, DNA CODING FOR SAME
;    TITLE OF INVENTION:  AND METHOD OF PRODUCING THE PROTEIN
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,558A
;      FILING DATE:  24-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JPA Hei 7-187135
;      FILING DATE:  24-JUL-1995
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  513 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      STRAIN:  MKN45
US-08-685-558A-18


  Query Match             20.0%;  Score 6;  DB 4;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 EPLVLK 15
           ||||||
Db     177 EPLVLK 182


RESULT  14
US-08-277-231A-4
; Sequence 4, Application US/08277231A
; Patent No. 5643725
;  GENERAL INFORMATION:
;    APPLICANT:  Green, Bruce A.
;    APPLICANT:  Brinton, Charles C.
;    TITLE OF INVENTION:  Sequence and Analysis of LKP Pilin
; Patent No. 5643725
;    TITLE OF INVENTION:  Structural Genes and The LKP Pili Operon of No. 5643725typa
;    TITLE OF INVENTION:  Haemophilus Influenzae
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/277,231A
;      FILING DATE:  19-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carroll, Alice O.
;      REGISTRATION NUMBER:  32,542
;      REFERENCE/DOCKET NUMBER:  ACC94-02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  741 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-277-231A-4


  Query Match             20.0%;  Score 6;  DB 1;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 79;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 TARIGE 10
           ||||||
Db     157 TARIGE 162


RESULT  15
US-08-473-750-7
; Sequence 7, Application US/08473750
; Patent No. 5834187
; Patent No. 5834187 5786143
;  GENERAL INFORMATION:
;    APPLICANT:  Green, Bruce A.
;    APPLICANT:  Brinton, Jr., Charles C.
;    TITLE OF INVENTION:  Sequence and Analysis of LKP Pilin
; Patent No. 5834187
; Patent No. 5834187 5786143
;    TITLE OF INVENTION:  Structural Gene and the LKP Pili Operon of No. 5834187 5786
;    TITLE OF INVENTION:  Haemophilus Influenzae
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;       STREET:  Two Militia Drive
;       CITY:  Lexington
;       STATE:  Massachusetts
;       COUNTRY:  US
;       ZIP:  02173
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/473,750
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/277,321
;       FILING DATE:  19-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Carroll, Alice O.
;       REGISTRATION NUMBER:  33,542
;       REFERENCE/DOCKET NUMBER:  ACC94-02B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-861-6240
;       TELEFAX:  617-861-9540
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  741 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-473-750-7


  Query Match             20.0%;  Score 6;  DB 2;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 79;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 TARIGE 10
             ||||||
Db     157 TARIGE 162


Search completed: July 31, 2002, 15:04:07
Job time: 22 sec
```

L7 ANSWER 26 OF 36 MEDLINE
AN 1999182371 MEDLINE
DN 99182371 PubMed ID: 10082470
TI Activation of receptor for advanced glycation end products: a mechanism for chronic vascular dysfunction in diabetic vasculopathy and atherosclerosis.
AU Schmidt A M; Yan S D; Wautier J L; Stern D
CS Division of Surgical Science, Department of Surgery, College of Physicians & Surgeons of Columbia University, New York, NY 10032, USA.
SO CIRCULATION RESEARCH, (1999 Mar 19) 84 (5) 489-97. Ref: 89
Journal code: 0047103. ISSN: 0009-7330.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 199903
ED Entered STN: 19990402
Last Updated on STN: 19990402
Entered Medline: 19990324
AB Receptor for advanced glycation end products (**RAGE**) is a member of the immunoglobulin superfamily of cell surface molecules and engages diverse ligands relevant to distinct pathological processes. One class of **RAGE** ligands includes glycoxidation products, termed advanced glycation end products, which occur in diabetes, at sites of oxidant stress in tissues, and in renal failure and **amyloidoses**. **RAGE** also functions as a signal transduction receptor for **amyloid** beta peptide, known to accumulate in Alzheimer disease in both affected brain parenchyma and cerebral vasculature. Interaction of **RAGE** with these ligands enhances receptor expression and initiates a positive feedback loop whereby receptor occupancy triggers increased **RAGE** expression, thereby perpetuating another wave of cellular activation. Sustained expression of **RAGE** by critical target cells, including endothelium, smooth muscle cells, mononuclear phagocytes, and neurons, in proximity to these ligands, sets the stage for chronic cellular activation and tissue damage. In a model of accelerated atherosclerosis associated with diabetes in genetically manipulated mice, blockade of cell surface **RAGE** by infusion of a soluble, truncated form of the receptor completely suppressed enhanced formation of vascular lesions. Amelioration of atherosclerosis in these diabetic/atherosclerotic animals by soluble **RAGE** occurred in the absence of changes in plasma lipids or glycemia, emphasizing the contribution of a lipid- and glycemia-independent mechanism(s) to atherogenesis, which we postulate to be interaction of **RAGE** with its ligands. Future studies using mice in which **RAGE** expression has been genetically manipulated and with selective low molecular weight **RAGE** inhibitors will be required to definitively assign a critical role for **RAGE** activation in diabetic vasculopathy. However, sustained receptor expression in a microenvironment with a plethora of ligand makes possible prolonged receptor stimulation, suggesting that interaction of cellular **RAGE** with its ligands could be a factor contributing to a range of important chronic disorders.